Regression-based Testing for Generalized Cyclical Fractional Unit Roots*

Uwe Hassler[a], Paulo M. M. Rodrigues[b] and Antonio Rubia[c]

[a] Goethe University Frankfurt

[b] University of Algarve

[c] University of Alicante

September 2007

Abstract

In this paper, we propose a least-squares based procedure intended to test for cyclical and/or seasonal integration against fractional alternatives in the time domain. This approach belongs to the Lagrange-multiplier framework for long-memory series studied in Robinson (1991, 1994), Agliakloglou and Newbold (1994), Agliakloglou, Newbold and Wohar (1993), Tanaka (1999), Hassler and Breitung (2002, 2006) and Nielsen (2004). The family of regression-based tests we discuss has convenient methodological advantages. It can be easily implemented in practical settings, it is flexible enough to account for a broad family of long- and short-memory specifications, it has power against different type of alternative hypotheses, and it allows inference to be conducted under critical values which are drawn from a standard chi-squared distribution, independently of the long-memory parameters.

*This version: September 2007. Please do not quote without permission from the authors.

1 Introduction

Modelling and forecasting macroeconomic cycles as well as financial variables is at the forefront of the applied time-series econometrics literature. In this paper, we propose a least-squares based procedure intended to test for cyclical and/or seasonal integration against fractional alternatives in the time domain. This approach belongs to the Lagrange-multiplier framework for long-memory series studied in Robinson (1991, 1994), Agliakloglou and Newbold (1994), Agliakloglou, Newbold and Wohar (1993), Tanaka (1999), Hassler and Breitung (2002, 2006) and Nielsen (2004). The family of regression-based tests we discuss has convenient methodological advantages. It can be easily implemented in practical settings, it is flexible enough to account for a broad family of long- and short-memory specifications, it has power against different type of alternative hypotheses, and it allows inference to be conducted under critical values which are drawn from a standard chi-squared distribution, independently of the long-memory parameters.

More specifically, we discuss a test formally intended to detect cyclical long memory patterns embedded in the autoregressive filter

$$(1-L)^{d_0}\,(1+L)^{d_\pi}\prod_{v=1}^{k}(1-2\eta_v L+L^2)^{d_v}$$

where d_l are possibly non-integer values, η_v characterize the cyclical behavior (periodicity) of the data, and L is the conventional back-shift operator. The filter is able to capture both long-range dependence and periodic cyclical fluctuations through the convolution of Gegenbauer processes. These generate theoretical autocovariances that decay hyperbolically and sinusoidally, a feature that is manifested in a number of periodic time series. Particular cases of this general setting include pure cyclical and seasonal models, which are routinely applied to fit both economic and non-economic variables. For instance, cyclical models have been used to explain macroeconomic dynamics by Gray *et al.* (1989, 1994), Ramachandran and Beaumont (2001), Barkoulas *et al.* (2001), Gil-Alaña (2001, 2004), Caporale and Gil-Alaña (2006) and Smallwood and Norrbin (2006), among others. Recent studies focusing on non-economic variables have analyzed, for instance, atmospheric levels of CO_2 (Woodward *et al.*, 1998), wind speed (Bouette *et al.*, 2006), or power demand (Soares and Souza 2006). The literature related to the seasonal models embedded in this general framework (*e.g.*, both integrated and fractionally integrated seasonal models) is overwhelming.

The regression-based testing procedure we propose in this paper is intended to provide a formal tool for pretesting hypothesis about the extent of cyclic and non-cyclic persistence in these areas.

The remaining of the paper is organized as follows. Section 2 introduces the general setting and discusses the set of sufficient conditions for the tests. Section 3 discusses the specific form of the regression to be used as well as the relevant test statistics. The asymptotic distribution is discussed in several theorems. Section 4 analyzes the finite-sample performance of the tests by means of Monte Carlo experimentation. Section 5 summarizes the main conclusions. Finally, the mathematical proofs of the main results are collected in a technical appendix.

In what follows, $\Rightarrow$, $\xrightarrow{p}$, $\rightarrow$ denotes weak convergence in distribution, convergence in probability, and convergence of a series of real numbers, respectively, as the number of observations is allowed to diverge. The conventional notation $o(1)$ $(o_p(1))$ is used to represent a series of numbers (random numbers) converging to zero (in probability), while $O(1)$, $(O_p(1))$ denotes a series of numbers (random numbers) that are bounded (in probability). The notation $[x]$ is used for the integer value of the real-valued number x, and $\mathbb{I}_{(\cdot)}$ is the indicator function that takes value equal to one if the condition in the subscript is fulfilled and zero otherwise. Finally, vectors and matrices are denoted through bold letters.

2 The general fractionally integrated model

Let $\xi_\lambda(L;\delta)$ be a Gegenbauer polynomial in the lag operator defined as follows,

$$\xi_\lambda(L;\delta) = (1 - 2\cos\lambda L + L^2)^\delta \tag{1}$$

where the long-memory parameter δ can take non-integer values and controls the extent of time-series dependence. The parameter λ is a so-called Gegenbauer frequency in $[0,\pi]$, and controls the periodicity of the resulting time series.

Define the following generalization of (1), given the set of long-memory parameters $\boldsymbol{\delta} = (\delta_0, , \delta_1, ..., \delta_k, \delta_\pi)'$, and the vector $\boldsymbol{\lambda} = (\lambda_1, ..., \lambda_k)$,

$$\Delta_{\boldsymbol{\lambda}}(L;\boldsymbol{\delta}) \equiv (1-L)^{\delta_0}(1+L)^{\delta_\pi}\prod_{v=1}^{k}\xi_{\lambda_v}(L;\delta_v) \tag{2}$$

such that $0 < \lambda_1 < \cdots < \lambda_k < \pi$, and $k \geq 1$. Denote $\boldsymbol{\gamma} = (0, \lambda_1, .., \lambda_k, \pi)'$ as the vector collecting all the frequencies involved, whose elements we denote as γ_s, $s = 1, ..., k+2$. The resultant filter allows for multiple cyclical components on the seasonal frequencies γ_s, $s > 1$, as well as a long-run trend at the zero frequency. Furthermore, this model encompasses quite different types of models. Major examples for empirical purposes include pure cyclical models and fractionally integrated unit-root models (which arise by restricting $\boldsymbol{\delta}$), and pure seasonal models (which arise by restricting $\boldsymbol{\lambda}$). We shall briefly discuss the properties of these restricted models at the end of this section.

We consider that the observable process, $\{x_t,\ t = 1, ..., T\}$, admits the following characterization under the null hypothesis

$$\bar{\Delta}_{\boldsymbol{\lambda}}(L;\boldsymbol{\delta})\, x_t = \varepsilon_t \tag{3}$$

where $\bar{\Delta}_{\boldsymbol{\lambda}}(L;\boldsymbol{\delta})$ is a possibly restricted version of $\Delta_{\boldsymbol{\lambda}}(L;\boldsymbol{\delta})$ in the terms discussed previously, and ε_t is a covariance stationary noise process with spectral density that is bounded and bounded away from zero at all frequencies. In the most general case, $\bar{\Delta}_{\boldsymbol{\lambda}}(L;\boldsymbol{\delta}) = \Delta_{\boldsymbol{\lambda}}(L;\boldsymbol{\delta})$, we will say that x_t is generated by a *General Fractionally Integrated* process of order $\boldsymbol{\delta}$, denoted as x_t ~GFI($\boldsymbol{\delta}$).

For empirical purposes, the main interest lies in testing whether $\boldsymbol{\delta} = \mathbf{d}$, with $\mathbf{d} \in R^n$, $1 \leq n \leq k+2$, is specified *a priori*, against the alternative for which the order of integration is $\mathbf{d}^*$, $\mathbf{d}^* = \mathbf{d} + \boldsymbol{\theta}$, $\boldsymbol{\theta} \neq \mathbf{0}$. Thus, the hypothesis of interest is generally stated as

$$\mathrm{H}_0 : \boldsymbol{\theta} = \mathbf{0}, \tag{4}$$

against $\mathrm{H}_1 : \boldsymbol{\theta} \neq \mathbf{0}$. The set of assumptions that we shall consider throughout the paper (assumption $\mathcal{A}$) is presented and discussed below.

Assumption $\mathcal{A}$:

i) The observable process $\{x_t,\ t = 1, ..., T\}$ is generated by $\Delta_{\boldsymbol{\gamma}}(L; \mathbf{d})\, x_t = \varepsilon_t \mathbb{I}_{(t>0)}$, with $\Delta_{\boldsymbol{\gamma}}(L; \mathbf{d})$ defined in (2), $\mathbf{d}$ being a possibly non-integer vector in R^m, $m \equiv k+2$, and $\boldsymbol{\gamma} = \in [0, \pi]$.
ii) $\varepsilon_t \sim iid\,(0, \sigma^2)$, and $E\left(\varepsilon_t^4\right) = \kappa < \infty$.

Some comments follow. We consider the most general case under the null hypothesis, which is given by x_t ~GFI($\mathbf{d}$). Simpler specifications (*e.g.*, pure seasonal models) arise considering restricted versions of $\Delta_{\boldsymbol{\lambda}}(L; \mathbf{d})\, x_t$, i.e., assuming $\bar{\Delta}_{\boldsymbol{\lambda}}(L; \mathbf{d})\, x_t = \varepsilon_t$ with $\mathbf{d} \in R^{m^*}$, $1 \leq m^* < k+2$. Condition *i*) also implies that $x_j = \varepsilon_j = 0$ for any $j \leq 0$, *i.e.*, we consider the realizations from a truncated stochastic process; see Marinucci and Robinson (1999), Robinson (2004) for a discussion. The i.i.d. assumption in *ii*) is stronger than necessary and can considerably be weakened by the martingale-difference sequence (MDS) hypothesis which allows, for instance, for time-varying conditional variance patterns, and serves as a basis for short-run dynamics. We shall analyze this possibility later on. Finally, we do not require normality, since this is not essential to derive the asymptotic theory, but we note that efficiency of Gaussian-score based procedures would only be attainable under that restriction.

2.1 Restricted long-memory models

It is worth referring to some leading models nested in (2). First, pure cyclical models arise by setting $\delta_0 = \delta_\pi = 0$, which leads to

$$\bar{\Delta}_{\boldsymbol{\lambda}}(L; \boldsymbol{\delta}) \equiv \prod_{v=1}^{k} \xi_{\lambda_v}(L; \delta_v)\,. \tag{5}$$

When $k = 1$, x_t is said to be generated by a GARMA model, whereas $k > 1$ leads to the so-called k-factor GARMA models; see Gray *et al.* (1996) and Ramachandran and Beaumont (2001) for a discussion of the statistical properties of these models. Further generalizations (for instance, allowing for stationary short-run dynamics) are able to encompass both ARMA and ARFIMA models as particular cases.

Pure seasonal models arise by restricting both the dimension and the value of $\boldsymbol{\lambda}$ aiming to relate its frequencies to the periodicity of the data, say S. For instance, if S is even, then $k = [S/2] - 1$, $\lambda_v = 2\pi v/S$, $v = 1, ..., k$, thereby leading to

$$\bar{\Delta}_{\boldsymbol{\lambda}}(L; \boldsymbol{\delta}) \equiv (1-L)^{\delta_0}(1+L)^{\delta_\pi} \prod_{v=1}^{[S/2]-1} \xi_{\lambda_v}(L; \delta_v) \tag{6}$$

with $\boldsymbol{\delta} \in R^{k+2}$. When S is odd, the component $(1+L)^{d_\pi}$, which corresponds to a cycle of two periods, is simply omitted and the model has $[S/2]$ parameters. A special case is $\bar{\Delta}_{\boldsymbol{\lambda}}(L; \mathbf{1}) = (1 - L^S)$, *i.e.*, the well-known seasonal random-walk filter. By allowing non-integer values in $\boldsymbol{\delta}$, x_t is said to be generated by a seasonal fractionally integrated process of order $\boldsymbol{\delta}$; see, among others, Hassler (1994) and references therein.

Finally, the well-known fractional unit root model also arises after removing all the terms related to the non-zero frequencies, *i.e.*, by considering $\bar{\Delta}_{\boldsymbol{\lambda}}(L; d_0) = (1-L)^{\delta_0}$.

3 Testing procedures: asymptotic analysis

We start our analysis by introducing notation and defining some important variables in this context. The first definition introduces the weighting scheme that, for any frequency $\alpha \in [0, \pi]$, allows us to construct partial sum processes that convey statistical information about the order of integration at the α frequency under the alternative. The second and third definitions introduce the way these processes are computed, and some relevant vector notation, respectively.

Definition 3.1. *For all $j \geq 1$ and $\alpha \in [0, \pi]$, define the weighting variable $\omega_j(\alpha)$ as follows,*

$$\omega_j(\alpha) = \begin{cases} 1/j, & if\ \alpha = 0 \\ 2j^{-1}\cos(j\alpha), & if\ \alpha \in (0, \pi) \\ (-1)^j/j, & if\ \alpha = \pi \end{cases}. \tag{7}$$

Definition 3.2. *Given the real-valued stochastic process $\{x_t, t \geq 1\}$, the filtered series, $\widehat{\varepsilon}_t$, which obtains under $H_0 : x_t \sim GFI(\mathbf{d})$, is determined as*

$$\widehat{\varepsilon}_t = \Delta_{\boldsymbol{\lambda}}(L; \mathbf{d})\, x_t. \tag{8}$$

Also, *given $\{\widehat{\varepsilon}_t, t \geq 1\}$ and some $\alpha \in [0, \pi]$, define the (truncated) partial sum process $\widehat{\varepsilon}^*_{\alpha,t-1}$ as*

$$\widehat{\varepsilon}^*_{\alpha,t-1} = \sum_{j=1}^{t-1} \omega_j(\alpha)\, \widehat{\varepsilon}_{t-j}. \tag{9}$$

Definition 3.3. *Given* $\boldsymbol{\gamma} \equiv (0, \lambda_1, ..., \lambda_k, \pi)'$, *define* $\widehat{\boldsymbol{\varepsilon}}^*_{\boldsymbol{\gamma},t-1} = \left(\widehat{\varepsilon}^*_{\gamma_1,t-1}, ..., \widehat{\varepsilon}^*_{\gamma_{k+2},t-1}\right)'$, *i.e.,* $\widehat{\boldsymbol{\varepsilon}}^*_{\boldsymbol{\gamma},t-1} = \sum_{j=1}^{t-1} \boldsymbol{\omega}_j(\boldsymbol{\gamma}) \widehat{\varepsilon}_{t-j}$, *with* $\boldsymbol{\omega}_j(\boldsymbol{\gamma}) = \left(\omega_j(\gamma_1), ..., \omega_j(\gamma_{k+2})\right)'$ and where each element is as defined in Definition 3.1.

Following Robinson (1991, 1994b), Agiakloglou and Newbold (1994), Tanaka (1999), and Hassler and Breitung (2002, 2006), we can define a general testing strategy based on the score principle which allows us to carry out individual or joint inference in the general fractionally integrated context. In particular, assume that the researcher wants to test a hypothesis involving n, $1 \leq n \leq m$, long-memory parameters of the model $\Delta_{\boldsymbol{\gamma}}(L; \mathbf{d}) x_t = \varepsilon_t$. Without loss of generality, and only for simplicity of notation, we assume that these correspond to the first n elements in $\mathbf{d}$, related to the first n frequencies in $\boldsymbol{\gamma}$. Recall that the null hypothesis $\mathrm{H}_0 : \boldsymbol{\theta} = \mathbf{0}$, and that the alternative hypothesis implies the partition $\boldsymbol{\theta} = \left(\boldsymbol{\theta}'^*_n, \mathbf{0}'_{m-n}\right)'$, with $\boldsymbol{\theta}^*_n = (\tau_1, ..., \tau_n)'$, and some $\tau_s \neq 0$.

Proposition 3.1. *Given* $\{x_t, t = 1, ..., T\}$, *the null hypothesis* $\mathrm{H}_0 : x_t \sim$GFI$(\mathbf{d})$, *for* $\mathbf{d} \in R^m$, *can be tested against the alternative* $\mathrm{H}_1 : x_t \sim$GFI$(\mathbf{d} + \boldsymbol{\theta})$, *with* $\boldsymbol{\theta} = (\boldsymbol{\theta}^*_n, \mathbf{0}_{m-n})'$, $\boldsymbol{\theta} \neq \mathbf{0}$, *and* $1 \leq n \leq m$, *through the squares of the t-statistic for the joint significance of the estimated* $\{\phi_s\}_{s=1}^n$ *parameters, say* $\Upsilon^{(n)}$, *in the auxiliary regression:*

$$\widehat{\varepsilon}_t = \sum_{s=1}^{n} \phi_s \widehat{\varepsilon}^*_{\gamma_s,t-1} + e_t \tag{10}$$

i.e.,

$$\Upsilon^{(n)} = \left[\sum_{t=2}^{T} \hat{\varepsilon}_t \left(\widehat{\boldsymbol{\varepsilon}}^*_{\boldsymbol{\gamma},t-1}\right)\right]' \left[\widehat{\sigma}^2_e \sum_{t=2}^{T} \left(\widehat{\boldsymbol{\varepsilon}}^*_{\boldsymbol{\gamma},t-1} \widehat{\boldsymbol{\varepsilon}}^{*\prime}_{\boldsymbol{\gamma},t-1}\right)\right]^{-1} \left[\sum_{t=2}^{T} \hat{\varepsilon}_t \left(\widehat{\boldsymbol{\varepsilon}}^*_{\boldsymbol{\gamma},t-1}\right)\right], \tag{11}$$

with $\widehat{\sigma}^2_e$ *being the LS estimate of the variance of the residuals* e_t, *and* $\widehat{\varepsilon}_t$ *and* $\widehat{\boldsymbol{\varepsilon}}^*_{\boldsymbol{\gamma},t-1}$, *as given in Definitions 3.2 and 3.3, respectively.*

Theorem 3.1. *Let Assumption* $\mathcal{A}$ *hold true, and let* $\Upsilon^{(n)}$ *be the test statistic defined previously. Under* $\mathrm{H}_0 : \boldsymbol{\theta} = 0$, *and as* $T \to \infty$, *it follows that*

$$\Upsilon^{(n)} \Rightarrow \chi^2_{(n)}$$

where $\chi^2_{(n)}$ *stands for a Chi-squared distribution with n degrees of freedom.*

Corollary 3.1. *Individual inference on the long-memory parameter related to the s-th frequency,* $1 \leq s \leq m$, *is embedded in Proposition 3.1. Under the alternative,* $\boldsymbol{\theta}$ *is now restricted* *to take value $\tau \neq 0$ in its s-th entry, and zero otherwise. In particular, the auxiliary regression reduces to*

$$\widehat{\varepsilon}_t = \phi_s \, \widehat{\varepsilon}^*_{\gamma_s,t-1} + e_{s,t},$$

and the relevant test statistic, say Υ_{γ_s}, is the squared t-statistic for the significance of the estimate of ϕ_s, i.e.,

$$\Upsilon_{\gamma_s} = \left(\sum_{t=2}^{T} \widehat{\varepsilon}_t \widehat{\varepsilon}^*_{\gamma_s,t-1} \right)^2 \left(\widehat{\sigma}^2_{s,e} \sum_{t=2}^{T} \left(\widehat{\varepsilon}^*_{\gamma_s,t-1} \right)^2 \right)^{-1} . \tag{12}$$

Corollary 3.2. *Similarly, it could be of interest to test for the restricted joint hypothesis $\boldsymbol{\theta}^*_n = (\tau, ..., \tau)' = \tau \mathbf{1}'_n$, $\tau \neq 0$. This is the case, for instance, when analyzing the suitability of so-called rigid models, which assume homogeneity in the order of fractional integration, see Porter-Hudak (1990) and Hassler (1994). Since this implies $\phi_l = \phi$, $l = 1, ..., m$, the auxiliary regression reduces to*

$$\widehat{\varepsilon}_t = \phi \, \widehat{\varepsilon}^*_{\boldsymbol{\gamma},t-1} + e_t,$$

with $\widehat{\varepsilon}^*_{\boldsymbol{\gamma},t-1} \equiv \sum_{l=1}^{n} \widehat{\varepsilon}^*_{\gamma_l,t-1}$, and the relevant statistic, say $\bar{\Upsilon}^{(n)}$, is now given by

$$\bar{\Upsilon}^{(n)} = \left(\sum_{t=2}^{T} \widehat{\varepsilon}_t \widehat{\varepsilon}^*_{\boldsymbol{\gamma},t-1} \right)^2 \left(\widehat{\sigma}^2_e \sum_{t=2}^{T} \left(\widehat{\varepsilon}^*_{\boldsymbol{\gamma},t-1} \right)^2 \right)^{-1} . \tag{13}$$

Then, under Assumption $\mathcal{A}$, and as $T \to \infty$, it holds that

$$\begin{aligned} \Upsilon_{\phi,s} &\Rightarrow \chi^2_{(1)} \\ \bar{\Upsilon}^{(n)} &\Rightarrow \chi^2_{(1)} \end{aligned}$$

because only one restriction is implied.

Remark 3.1. The tests are asymptotically equivalent to the frequency domain LM tests studied in Robinson (1994), and the time domain LM test considered in Tanaka (1999) –restricted to the fractional unit root model. In our case, the Fisher Information matrix is estimated as the outer product of gradients. The tests are also asymptotically equivalent to the general likelihood-based tests in Nielsen (2004), discussed in the context of maximum-likelihood model estimation. The LM regression-based test in Breitung and Hassler (2002), focused on the fractional unit root model, $\bar{\Delta}_{\boldsymbol{\lambda}} (L; \boldsymbol{\delta}) = (1 - L)^{d_0}$, arises as a restricted case in our context; see also Nielsen (2005), Hassler and Breitung (2006), and Demetrescu,Kuzin and Hassler (2007). It is worth mentioning that, as remarked in Nielsen (2004), the experimental simulations in Tanaka (1999), and Breitung and Hassler (2002), show that in finite samples the time domain fractionally-integrated unit-root tests tend to be superior to the frequency domain test, both in size and power behavior, so a similar performance is likely to be observed in a more general setting as well.

Remark 3.3. As discussed in Breitung and Hassler (2002), the auxiliary regression centered on the zero frequency, $\widehat{\varepsilon}_t = \phi_0 \, \widehat{\varepsilon}^*_{0,t-1} + e_t$, is reminiscent of the Dickey-Fuller regression and the Wald-test in Dolado, Gonzalo and Mayoral (2002). Meaningful difference arise, nevertheless, since in the DF test the regressor is FI(0) under the alternative, whereas $\widehat{\varepsilon}^*_{\gamma_1,t-1}$ is FI$(d+\tau)$ owing to the different types of weights used in constructing these variables. Similarly, for pure seasonal models, the general auxiliary regression in Proposition 3.1. is reminiscent of the HEGY regression, in the sense that the regressors $\widehat{\varepsilon}^*_{\gamma_s,t-1}$ are weighted linear combinations of lags of $\hat{\varepsilon}_t$ related to a specific frequency. Further differences arise in this case, because regressors in the HEGY context are ensured to be mutually orthogonal by construction, whereas the LM regressors in (10) are not. This feature warns against estimating the auxiliary regression (10) with n parameters, and then testing a hypothesis on the basis of this which involves a smaller subset of n_1 parameters (as it is nevertheless possible in the HEGY test).

Remark 3.5. We assume that the vector of frequencies $\boldsymbol{\gamma}$ is known. Indeed, this is the case for pure seasonal models, but in general terms it may be restrictive when analyzing cyclical models by means of Gegenbauer polynomials. Several approaches with high rate of convergence have been proposed to estimate consistently Gegenbauer-frequencies in the semi-parametric literature; see, among others, Yajima (1996), Giriatis, Hidalgo, and Robinson (2001), Hidalgo and Soulier (2004), and Hidalgo (2005). In any case, when using sample estimates for subsequent inference, it should be noticed that the performance of the test statistics may be subject to the potential distortions that often arises as a result of (small-sample) biases when inferring the unknown elements of $\boldsymbol{\gamma}$.

Example: Consider the pure seasonal quarterly case to illustrate the general testing principle we have discussed. Assume that the interest lies in testing the suitability of the seasonal unit root model, $(1-L^4)\,x_t = \varepsilon_t$, against a more general case in which the order of seasonal integration is possibly a non-integer value $1+\tau$, $\tau \neq 0$, but believed to be common at all frequencies, *i.e.*, $(1-L^4)^{d+\tau}\,x_t = \varepsilon_t$. Therefore, we have $\boldsymbol{\gamma} = (0, \pi/2, \pi)'$, $m = n = 3$, and the testing procedure for these rigid models is that described in Corollary 3.2. Therefore, we first compute $\{\widehat{\varepsilon}_t\}$ by differencing the series under the null hypothesis, thereby obtaining $\widehat{\varepsilon}_t = x_t - x_{t-4}$, and then compute the regressor $\widehat{\varepsilon}^*_{\boldsymbol{\gamma},t-1}$ as discussed, *i.e.*,

$$\begin{aligned}
\widehat{\varepsilon}^*_{\boldsymbol{\gamma},t-1} &= \widehat{\varepsilon}^*_{0,t-1} + \widehat{\varepsilon}^*_{\pi/2,t-1} + \widehat{\varepsilon}^*_{\pi,t-1} \\
&= \sum_{j=1}^{t-1} \left(\frac{1}{j} + \frac{(-1)^j}{j} + \frac{2\cos(j\pi/2)}{j} \right) \widehat{\varepsilon}_{t-j} = \sum_{j=1}^{t-1} \frac{\widehat{\varepsilon}_{t-4j}}{j}
\end{aligned}$$

with $\widehat{\varepsilon}_t = 0$ for all $t \leq 0$. Note that the resulting weighting scheme, $(j^{-1}L^4)$, corresponds to the expansion of $\log[(1-L^4)]$, which by construction ensures power against quarterly seasonal fractional integration.

3.1 Short-memory dynamics

Assumption $\mathcal{A}$ imposes the particularly restrictive condition that $\varepsilon_t \sim iid(0, \sigma^2)$. A more realistic approach allows for weakly dependent errors in the model. In particular, we require (say, **Assumption** $\mathcal{A}^*$) the observable time series $\{x_t, t = 1, ..., T\}$ to be a realization of the stochastic process:

$$\begin{aligned} \Delta_{\boldsymbol{\lambda}}(L; \mathbf{d})\, x_t &= \varepsilon_t \mathbb{I}_{(t>0)}, \quad t = 1, ..., T \\ \alpha(L)\, \varepsilon_t &= u_t \end{aligned} \tag{14}$$

where $\alpha(L) = 1 - \sum_j^p \alpha_j L^j$, $p \geq 0$, such that $\alpha(z)$ has all its roots outside of the unit circle, and $\{u_t\}$ is a MDS with respect to $\mathcal{F}_t$, $\mathcal{F}_t = \sigma(u_j : j \leq t)$ such that $E\left(|u_t|^{4+\nu}\right) < \infty$ for some $v > 0$.

More generally, for practical purposes, the short-run dynamics of the process may be characterized by any stationary and invertible general linear process, $\varepsilon_t = \sum_{s=0}^{\infty} b_j u_{t-j}$, under conditions which ensure that the AR(p) model is a good approximation for a finite p of the underlying AR representation of the true model. Demetrescu, Kuzin and Hassler (2007) discuss these conditions, which require only mild summability restrictions on the coefficients $\{b_j\}$, and the finiteness of certain high-order cumulants. The value of p may then be inferred through standard lag-selection methods.

We consider two alternative ways to deal with short-run dynamics. The first one is similar to Agiakloglou and Newbold (1994) and Breitung and Hassler (2002), and is based on the estimated residuals from an appropriate autoregression. The second approach is based on the same strategy as in Demetrescu *et al.* (2007), and considers inference in a suitably augmented regression framework.

Proposition 3.2. *Given* $\widehat{\varepsilon}_t = \Delta_{\boldsymbol{\lambda}}(L; \mathbf{d})\, x_t$, *let* $\{\tilde{u}_t\}$ *be the least-squares residuals from the p-th order autoregression:*

$$\widehat{\varepsilon}_t = \alpha_1 \widehat{\varepsilon}_{t-1} + ... + \alpha_p \hat{\varepsilon}_{t-p} + \tilde{u}_t, \qquad t = p+1, ..., T. \tag{15}$$

Then, the null hypothesis can be tested, under Assumption $\mathcal{A}^*$, *through the squared statistic for the joint-significance of the* $\{\phi_l^*\}_{l=1}^{k+2}$ *estimates,* $l = 1 ..., n$, *in the following regression:*

$$\tilde{u}_t = \sum_{l=1}^{n} \phi_l^* \widehat{u}^*_{\gamma_s, t-1} + \sum_{k=1}^{p} \gamma_k \hat{\varepsilon}_{t-k} + e_t \tag{16}$$

where

$$\widehat{u}^*_{\gamma_s, t-1} \equiv \sum_{j=1}^{t-p-1} \omega_j(\gamma_s)\, \tilde{u}_{t-j}. \tag{17}$$

Proposition 3.3. *Alternatively, the p-th order augmented regression in Proposition 3.1., i.e.,*

$$\widehat{\varepsilon}_t = \sum_{l=1}^{n} \phi_l \widehat{\varepsilon}^*_{\gamma_s,t-1} + \left(\sum_{k=1}^{p} \zeta_k \widehat{\varepsilon}_{t-k} \right) + e_t, \tag{18}$$

ensures asymptotic invariance against stationary short-run dynamics in the the (squared) test statistic for the joint significance of $\{\phi_l^*\}_{l=1}^{k+2}$. *The test statistic uses in this case White's correction to estimate the asymptotic covariance matrix.*

Theorem 3.2. *Let* $\Upsilon_a^{*(n)}$ *and* $\Upsilon_b^{*(n)}$ *be the squared test statistics as described in Proposition 3.4 and 3.5. Then, under Assumption* $\mathcal{A}^*$, $H_0 : \boldsymbol{\theta} = 0$, *and as* T *is allowed to diverge, it follows that:*

$$\begin{aligned} \Upsilon_a^{*(n)} &\Rightarrow \chi^2_{(n)} \\ \Upsilon_b^{*(n)} &\Rightarrow \chi^2_{(n)} \end{aligned}$$

Proof. *See appendix.*

Remark 3.6. Demetrescu *et al.* (2007) analyze the performance of data-dependent methods to determine the order of augmentation, p, in finite samples. The rule of thumb proposed by Schwert (1989), which sets $p = \left[c(T/100)^{1/4}\right]$, where c is a positive constant, shows a relatively good performance in finite-samples over several alternative approaches.

Remark 3.7. Note that we have focused on the model $\Delta_{\boldsymbol{\lambda}}(L;\mathbf{d})(x_t - \mu_t) = \varepsilon_t \mathbb{I}_{(t>0)}$, by allowing different dynamics in ε_t, and restricting $\mu_t = 0$. As commented in Breitung and Hassler (2002), the simplest way to deal with non-zero deterministic patterns, $\mu_t \neq 0$, is to detrend x_t prior to computing the relevant test statistics. This does not affect the limit distributions of the relevant statistics; see the discussion in Robinson (1994b).

4 Finite-sample analysis

In this section we address the empirical properties of the regression-based LM test statistic in finite samples. The case for the fractionally-integrated unit root process, $\bar{\Delta}_{\boldsymbol{\lambda}}(L;\boldsymbol{\delta}) = (1-L)^{d_0}$, has been considered previously in Breitung and Hassler (2002) and Nielsen (2004), showing the good performance of the LM test, both in absolute terms and in relation to alternative procedures. We therefore focus on cyclical and seasonal models through different specifications of $\bar{\Delta}_{\boldsymbol{\lambda}}(L;\boldsymbol{\delta})$, as our experiment study contributes to the better understanding of the LM test out of the fractionally-integrated unit root context. We conduct three Monte Carlo experiments.

First, we first consider the simple pure cyclic model,

$$(1 - 2\cos\lambda_s L + L^2)^{d+\theta} x_t = \varepsilon_t$$

in order to analyze the empirical size and power of Υ_{λ_s}, asymptotically distributed as $\chi_{(1)}$, when testing against $d = 1$ with θ in $[-0.3, 0.3]$. We consider 5000 replications with sample sizes $T = \{100, 250\}$, and $\varepsilon_t \sim iid\mathcal{N}(0,1)$. The Gegenbauer frequency is set $\lambda_s = s\pi/10$, with $s = 1, ..., 9$. The rejection frequencies for a nominal significance level of 5% are shown in Table 1.

[Insert Table 1 around here]

The test shows approximately correct sizes even in small samples, and only minor differences following no particular pattern arise across the frequencies λ_s considered. For non-zero vales of θ, we observe several interesting features in the empirical power functions. First, given λ_s and T, power tends to exhibit a symmetric U-shape figure around the $\pi/2$ frequency, which is more evident for small values of $|\theta|$. Hence, the larger the difference $|\lambda_s - \pi/2|$, $\lambda_s \in (0, \pi)$, the more powerful the testing procedure seems to become. The dependence on the frequency the test is related to is not surprising, since the variance of the regressor (and hence, the signal-to-noise ratio and, eventually, the power of the test) depends on the specific frequency λ, and more generally, on $\boldsymbol{\gamma}$; see Appendix B for further technical details. Furthermore, if we compare the evidence obtained for Υ_λ to the experimental results in Breitung and Hassler (2002) for the zero-frequency case [see Table 1, pp.176], the power observed on the latter frequency is approximately of the same order as that for $\lambda = \pi/2$. This suggest that, everything else equal, fractionally-integrated dynamics are generally much more easily detected in the cyclical framework than in the zero-frequency context. Dealing with the non-zero frequency also has other benefits in terms of power. For fixed T and λ_s, the power functions tend to be symmetric around $\theta = 0$, since only the size of $\theta - 0$, and not its sign, seems to drive the probability of rejection. This does not seem to be the case for the zero-frequency case analyzed in Breitung and Hassler (2002), where the LM test is prone to reject more easily if $\theta < 0$. Finally, power is largely enhanced for $T = 250$, thus showing the consistency of the testing procedure even for finite small samples.

Second, we consider a more general two-factor cyclical model given by,

$$(1 - 2\cos\lambda_1 L + L^2)^{d_1+\theta_1}(1 - 2\cos\lambda_2 L + L^2)^{d_2+\theta_2} x_t = \varepsilon_t.$$

We want to address the ability of unrestricted joint test $\Upsilon^{(2)}$, distributed asymptotically as $\chi_{(2)}$, as well as the joint restricted test $\bar{\Upsilon}^{(2)}$ and the individual tests Υ_{λ_1} and Υ_{λ_2}, distributed asymptotically $\chi_{(1)}$, to detect fractionally-integrated dynamics. As before, we set $d_1 = d_2 = 1$, and θ_1, θ_2 in $[-0.3, 0.3]$, considering 5000 replications with sample length $T = \{100, 250\}$, and $\varepsilon_t \sim iid\mathcal{N}(0,1)$. The joint test $\Upsilon^{(2)}$ is expected to reject the null if fractional integration is present in, at least, some of the frequencies involved, while the individual tests should only reject when fractional integration occurs at the frequency they are related to. The restricted joint test $\bar{\Upsilon}^{(2)}$ should be more efficient than $\Upsilon^{(2)}$ when the restriction $\theta_{\lambda_1} = \theta_{\lambda_2}$ is true, but it is expected to exhibit less comparative power to reject the false null otherwise.

In view of the previous experiment, we expect the power function to depend on the value of $\boldsymbol{\gamma} = (\lambda_1, \lambda_2)'$. We set $\lambda_1 = 0.15 \approx \pi/20$, corresponding to the estimated frequency of the business cycle by the NBER, and consider what seems to be the most unfavorable frequency for the tests, given by $\lambda_2 = \pi/2$, which also corresponds to one of the harmonics of the quarterly and monthly seasonality. For frequencies $\lambda \in (0, \pi)$ away from $\pi/2$, further simulations (not reported here) showed a much better statistical performance both in size and power. The rejection frequencies for a nominal significance level of 5% are shown in Table 2.

[Insert Table 2 around here]

Several interesting features emerge from this experiment. First, we comment the results for the individual tests $\Upsilon_{0.15}$ and $\Upsilon_{\pi/2}$. When $d_1 = 1$, and $d_2 = 1 + \theta_2$, both tests have approximately correct size when θ_2 is close to zero. However, when $|\theta_2|$ departures from the origin, $\Upsilon_{0.15}$ may show size departures with respect to the nominal size, which are particularly important when $\theta_2 > 0$. This is also true for the $\Upsilon_{\pi/2}$ test when $d_2 = 1$ and $d_1 = 1 + \theta_1$, now noting massive size distortions for large $\theta_1 > 0$. These small-sample distortions are originated in residual autocorrelation in $\widehat{\varepsilon}_t$, and can be considerably reduced (and even eliminated) by augmenting the auxiliary regression with $\bar{p}$ lags of the dependent variable. Table 1 shows, for $\bar{p} = 2$, that augmentation is effective to remove this small-sample distortion, particularly in the region $\theta > 0$ in which the effect was more pervasive. As usual, empirical size is corrected at the expenses of power reductions, which in our context can be large for the alternatives $\theta > 0$. Finally, it is interesting to remark that, the empirical size approaches the asymptotic nominal level, the power of the $\Upsilon_{\pi/2}$ test is slightly smaller than observed when the data generating process only includes one seasonal factor. A similar feature can be observed in the case of $\Upsilon_{0.15}$.

In relation to the joint test statistics $\bar{\Upsilon}^{(2)}$ and $\Upsilon^{(2)}$, we observe that the restricted test is more powerful than the latter when the restriction $\theta_1 = \theta_2$ is true, but it is also considerably less efficient in the general context $\theta_1 \neq \theta_2$, particularly for small values of $|\theta|$. Both tests tend to reject more easily the (false) null when fractional integration is present at the frequency 0.15, *i.e.*, in the frequency for which the magnitude $|\lambda_s - \pi/2\ |$ is larger. For instance, if $d_1 = 1 - 0.1$ and $d_2 = 1$, the power of $\bar{\Upsilon}^{(2)}$ and $\Upsilon^{(2)}$ is, approximately, 39.8% and 48.7%, respectively. In contrast, for $d_1 = 1$ and $d_2 = 1 - 0.1$, the power is only 8.2% and 16.1%. When both θ_1 and θ_2 move away from the origin, the power of the joint tests, particularly that of $\Upsilon^{(2)}$, largely increases. We note that the power of $\Upsilon^{(2)}$ seems to be symmetric for the set of frequencies considered, whereas $\bar{\Upsilon}^{(2)}$ tends to reject more easily when $\theta_1 > 0$ and $\theta_2 < 0$ than in the converse case. For instance, the power for $\theta_1 = 0.3$ and $\theta_2 = -0.3$ is almost 100%, whereas it is around 25% for $\theta_1 = -0.3$ and $\theta_2 = 0.3$. In both cases, the power of the unrestricted test $\Upsilon^{(2)}$ is almost 100%. Finally, and as in the case of the one-factor model simulation, considering a larger sample, $T = 250$, leads to a considerable enhancement in the statistical properties of all the tests. We do not present these results to save space, but these are available upon request.

The last experiment considers again the two-factor filter $\bar{\Delta}_{\boldsymbol{\lambda}}(L; \boldsymbol{\delta}) = (1 - 2\cos\lambda_1 L + L^2)^{d_1+\theta_1}(1 - 2\cos\lambda_2 L + L^2)^{d_2+\theta_2}$ now allowing for stationary and invertible ARMA(1,1) dynamics in the error term, *i.e.*, we consider the model

$$\begin{aligned}\bar{\Delta}_{\boldsymbol{\lambda}}(L;\boldsymbol{\delta})\,x_t &= u_t\\ (1-aL)\,u_t &= (1-bL)\,\varepsilon_t,\end{aligned}$$

under the restriction $|a| < 1$, $|b| < 1$. As in Demetrescu *et al.* (2007), we set $a = 0.5$ and $b = -0.5$ and consider the lag-selection rule $p = \left[4(T/100)^{1/4}\right]$ in the augmented auxiliary regressions. The rejections frequencies for the individual and joint tests are shown in Table 3.

[Insert Table 3 around here]

The general conclusions that arise for the weakly-dependent case are similar to those observed for the i.i.d, although we observe several quantitative changes. Augmentation proves able to control the empirical size correctly for all the tests, and only a small undersizing effect can be observed in our simulations. However, and as shown in previous literature, ensuring a correct empirical size against general ARMA dynamics through augmentation in particularly small samples, as the one considered here, comes usually to the expense of potentially large reductions in power in relation to the i.i.d. case. This pervasive effect has been widely documented in the unit-root literature, where augmentation of the Dickey-Fuller regression is probably the most widely used procedure in applied settings. In fact, the power of the individual and the joint tests shows figures similar in magnitude to those observed in Demetrescu *et al.* (2007) for the fractionally integrated unit root case. By sharp contrast to the unit-root case, fortunately, power improves considerably faster in the seasonal context as more observations are available. For instance, for the ARMA model assumed, the power of $\Upsilon^{(2)}$ is not larger than 39% in the range of $\boldsymbol{\theta}$ considered when only 100 observations are available, corresponding to $\boldsymbol{\theta}^* = (-0, 3, 0.3)'$. For a larger sample of $T = 500$,, everything else equal, power increases up to 98%. Similarly, $\bar{\Upsilon}^{(2)}$ has a peak of approximately 30% for $T = 100$ when $\theta_1 = \theta_2 = -0.3$, which dramatically increases up to 99% for a sample of 500 observations. Finally, $\Upsilon_{0.15}$ and $\Upsilon_{\pi/2}$ have power of 43% and 28% under $\boldsymbol{\theta}^*$ when $T = \mathit{100}$, respectively, whereas for $T = 500$ power reaches 95% and 86%, respectively.

5 Conclusion

In this paper, we have considered a regression-based LM test in the time-domain that allows us to test for fractionally-integrated patterns against integer integration in general cyclical models defined on a number of seasonal (and non-seasonal) frequencies. The tests involving single or multiple parameters can be computed from simple least-squares regressions, and are asymptotically equivalent to the frequency-domain LM test in Robinson (1994) and the likelihood-based tests in Nielsen (2004), for which the relevant critical values obtain from a χ^2 distribution with many degrees as the restrictions being tested, and with independence of the order of integration. Augmented versions of these tests are asymptotically robust against weakly-dependent errors following unknown patterns under quite general conditions, and exhibit good statistical performance in samples of moderate size. This makes the general regression-based LM testing strategy a valuable tool for preliminary data analysis in applied settings in which the validity of parsimonious, yet potentially restrictive hypothesis related to the order of integration of the raw data, may be formally addressed in an intuitive and straightforward way.

Appendix

A. Lagrange Multiplier

The Gaussian log-likelihood function for $\beta = (\mathbf{d}^*, \sigma^2)'$, conditional on $\boldsymbol{\lambda}$ and the set of information $\mathbf{x}_T = \{x_t, t = -\infty, ..., T\}$, with initial observations $x_t = 0$, $t \leq 0$, is given by

$$\mathcal{L}(\beta|\mathbf{x}_T) = -\frac{T}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{t=1}^{T}\left[\Delta_{\boldsymbol{\gamma}}(L;\mathbf{d})\, x_t\right]^2 \tag{A.1}$$

and partial derivative

$$\begin{aligned}
\left.\frac{\partial\mathcal{L}(\beta|\mathbf{x}_T)}{\partial\boldsymbol{\theta}}\right|_{\mathrm{H}_0:\theta=0} &= -\frac{1}{\sigma^2}\sum_{t=1}^{T}\varepsilon_t\left.\left(\frac{\partial\varepsilon_t}{\partial\theta}\right)\right|_{\mathrm{H}_0:\theta=0} \\
&= -\frac{1}{\sigma^2}\sum_{t=1}^{T}\varepsilon_t\log\left[\Delta_{\boldsymbol{\lambda}}(L;\mathbf{1})\right]\varepsilon_t \\
&= -\frac{1}{\sigma^2}\sum_{t=1}^{T}\varepsilon_t\left(\log\left[1-L\right]+\log\left[1+L\right]+\sum_{v=1}^{k}\log\left[\xi_{\lambda_v}(L;1)\right]\right)\varepsilon_t \\
&= \frac{1}{\sigma^2}\sum_{t=1}^{T}\varepsilon_t\sum_{l=1}^{k+2}\left\{\sum_{j=1}^{\infty}\omega_j(\gamma_l)\, L^j\varepsilon_t\right\}
\end{aligned} \tag{A.2}$$

with $\omega_j(\gamma_l)$, $l = 1, ..., k+2$ given in Definition 3.1. Under Assumption $\mathcal{A}$, this expression reduces to

$$\left.\frac{\partial\mathcal{L}(\beta|\mathbf{x}_T)}{\partial\boldsymbol{\theta}}\right|_{\mathrm{H}_0:\theta=0} = \frac{1}{\sigma^2}\sum_{t=1}^{T}\varepsilon_t\sum_{l=1}^{k+2}\left\{\sum_{j=1}^{t-1}\omega_j(\gamma_l)\,\varepsilon_{t-j}\right\} \tag{A.3}$$

and after recalling that $\varepsilon^*_{\gamma_l,t-1} = \sum_{j=1}^{t-1}\omega_j(\gamma_l)\,\varepsilon_{t-j}$ and $\varepsilon^*_{t-1} = \sum_{l=1}^{k+2}\varepsilon^*_{\gamma_l,t-1}$, we have that

$$\left.\frac{\partial\mathcal{L}(\beta|\mathbf{x}_T)}{\partial\theta}\right|_{\mathrm{H}_0:\theta=0} = \frac{1}{\sigma^2}\left(\sum_{t=1}^{T}\varepsilon_t\varepsilon^*_{t-1}\right). \tag{A.4}$$

which serves as the basis for the test statistics considered in the main text.

B. Technical proofs

Before proving Theorem 3.1 and 3.2., we first introduce in a definition the limit expressions which characterize the asymptotic variances and covariances of the partial sum processes, and discuss two useful lemmas.

Definition B.1. *For any* $\lambda \in [0, \pi]$, *let*

$$\psi(\lambda) = \lim_{T \to \infty} \sum_{j=1}^{T} \omega_j^2(\lambda).$$

Straightforward calculus shows $\psi(\lambda) = \pi^2/6$, if $\lambda = \{0, \pi\}$, and $\psi(\lambda) = 2\left(\pi^2/3 - \pi\lambda + \lambda^2\right)$, *otherwise. Similarly, given* $\lambda_n, \lambda_m \in [0, \pi]$, $\lambda_n \neq \lambda_m$, *let*

$$\psi(\lambda_n, \lambda_m) = \lim_{T \to \infty} \sum_{j=1}^{T} \omega_j(\lambda_n)\, \omega_j(\lambda_m)$$

Note that $|\psi(\lambda_n, \lambda_m)| < \infty$ *and, in particular,*

$$\psi(\lambda_n, \lambda_m) = \begin{cases} -\psi(\lambda_m)/2 & if\ \lambda_n = 0, \lambda_m = \pi \\ \left(\psi(\lambda_m) - \lambda_m^2\right)/2 & if\ \lambda_n = 0, \lambda_m \in (0, \pi) \\ \left(\lambda_m^2 - \psi(\lambda_m)\right)/4 & if\ \lambda_n = \pi, \lambda_m \in (0, \pi) \end{cases},$$

where if λ_n,$\lambda_m \in (0, \pi)$, *then*

$$\begin{aligned} \psi(\lambda_n, \lambda_m) &= 2\pi/3 - \pi(\lambda_n + \lambda_m + |\lambda_n - \lambda_m|) \\ &\quad + \left((\lambda_n + \lambda_m)^2 + (|\lambda_n - \lambda_m|)^2\right)/2. \end{aligned}$$

Definition B.2. *Given the vector of ordered frequencies* $\boldsymbol{\gamma} \equiv (0, \lambda_1, ..., \lambda_{k+2}, \pi)'$, let $\boldsymbol{\omega}_j(\boldsymbol{\gamma}) = (\omega_j(0), \omega_j(\lambda_1), ..., \omega_j(\pi))'$, *denote* $\boldsymbol{\Gamma}_{\boldsymbol{\gamma}} = \lim_{T \to \infty} \sum_{t=1}^{T} \boldsymbol{\omega}_j(\boldsymbol{\gamma})\, \boldsymbol{\omega}_j(\boldsymbol{\gamma})'$, *i.e.*,

$$\boldsymbol{\Gamma}_{\boldsymbol{\gamma}} = \begin{pmatrix} \psi(0) & \psi(0, \lambda_1) & ... & \psi(0, \pi) \\ \psi(0, \lambda_1) & \psi(\lambda_1) & ... & \psi(\lambda_1, \pi) \\ \vdots & \vdots & \ddots & \vdots \\ \psi(\pi, 0) & \psi(\pi, \lambda_1) & ... & \psi(\pi) \end{pmatrix},$$

whit $\boldsymbol{\Gamma}_{\boldsymbol{\gamma}} < \infty$ *being a symmetric definite-positive matrix for* $\boldsymbol{\gamma} \in [0, \pi]$ *which admits the Cholesky decomposition, say* $\boldsymbol{\Gamma}_{\boldsymbol{\gamma}} = \boldsymbol{\Lambda}_{\boldsymbol{\gamma}}' \boldsymbol{\Lambda}_{\boldsymbol{\gamma}}$.

Lemma B.1. *Given* $\widehat{\varepsilon}_t = \Delta_{\boldsymbol{\lambda}}(L;\mathbf{d})\, x_t$ *and* $\boldsymbol{\gamma} = (\gamma_1, .., \gamma_{k+2})'$, *consider* $\widehat{\boldsymbol{\varepsilon}}^*_{\boldsymbol{\gamma},t-1} = \left(\widehat{\varepsilon}^*_{\gamma_1,t-1}, ..., \widehat{\varepsilon}^*_{\gamma_{k+2},t-1}\right)'$ *as given in Definition 3.3. Then, as* $T \to \infty$, *and under Assumption* $\mathcal{A}$:

i) $\widehat{\varepsilon}_t = \varepsilon_t$.
ii) $E\left(\widehat{\boldsymbol{\varepsilon}}^*_{\boldsymbol{\gamma},T}\right) = 0$, and $E\left(\widehat{\boldsymbol{\varepsilon}}^*_{\boldsymbol{\gamma},T}\widehat{\boldsymbol{\varepsilon}}^{*\prime}_{\boldsymbol{\gamma},T}\right) = \sigma^2\boldsymbol{\Gamma}_{\boldsymbol{\gamma}}$.
iii) Denote $\mathbf{I}_m$ *the identity matrix in* R^m. *Then,*

$$\frac{1}{\sigma^2\sqrt{T}}\left(\boldsymbol{\Lambda}^{-1}_{\boldsymbol{\gamma}}\sum_{t=2}^{T}\widehat{\varepsilon}_t\left(\widehat{\boldsymbol{\varepsilon}}^*_{\boldsymbol{\gamma},t-1}\right)\right) \Rightarrow \mathcal{N}(\mathbf{0}, \mathbf{I}_m).$$

Lemma B.2. *Denote* $\boldsymbol{\Omega}_T = T^{-1}\sum_{t=2}^{T}\left(\widehat{\boldsymbol{\varepsilon}}^*_{\boldsymbol{\gamma},t-1}\widehat{\boldsymbol{\varepsilon}}^{*\prime}_{\boldsymbol{\gamma},t-1}\right)$ *and let* $\widehat{\sigma}^2_e$ *the variance from the least-squares residuals. Then:*

i) $\boldsymbol{\Omega}_T \xrightarrow{p} \sigma^2\boldsymbol{\Gamma}_{\boldsymbol{\gamma}}$.
ii) $\widehat{\sigma}^2_e \xrightarrow{p} \sigma^2$.

Proof of Lemma B.1.
Part *i)* is obvious under Assumption $\mathcal{A}$, while the proof of *ii)* is immediate by taking expectations and noting $E(\varepsilon_t\varepsilon_j) = \sigma^2$ for all $t = j$, and zero otherwise. Part *iii)* states a multivariate invariance principle for the partial sum process $\mathbf{S}_{\boldsymbol{\gamma},T} \equiv \sum_{t=2}^{T}\widehat{\varepsilon}_t\widehat{\boldsymbol{\varepsilon}}^*_{\boldsymbol{\gamma},t-1}$. We use the short-notation $\boldsymbol{\omega}_j = \boldsymbol{\omega}_j(\boldsymbol{\gamma})$ and note that

$$\begin{aligned}\sum_{t=2}^{T}\widehat{\varepsilon}_t\widehat{\boldsymbol{\varepsilon}}^*_{\boldsymbol{\gamma},t-1} &= \sum_{t=2}^{T}\varepsilon_t\left(\sum_{j=1}^{t-1}\boldsymbol{\omega}_j\varepsilon_{t-j}\right) = \sum_{j=1}^{T-1}\boldsymbol{\omega}_j\left\{\sum_{t=j+1}^{T}\varepsilon_t\varepsilon_{t-j}\right\} \\ &= \sum_{j=1}^{T-1}\boldsymbol{\omega}_j\{Z_j\} = \sum_{j=1}^{T-1}\mathbf{Z}^*_{\gamma,j}, \text{ say.}\end{aligned}$$

The m-dimensional sequence $\left\{\mathbf{Z}^*_{\gamma,j}, \mathcal{G}_j\right\}$, with $\mathcal{G}_j = \sigma(Z_l : l \le j)$, is a MDS, because $E\left(\mathbf{Z}^*_{\gamma,j}\right) = \mathbf{0}$ and $\mathrm{E}\big(\mathbf{Z}^*_{\gamma,j}|\mathcal{G}_{j-1}\big) = \mathbf{0}$ for all j. Then, as T is allowed to diverge, the desired convergence follows directly from the Central Limit Theorem (CLT) for MDS (Hall and Heyde 1980). This in turns requires verifying some conditions, namely, $(\mathcal{C}_1)$ $T^{-1}\left(\mathbf{S}_{\boldsymbol{\gamma},T}\mathbf{S}'_{\boldsymbol{\gamma},T}\right) \xrightarrow{p} \sigma^4\boldsymbol{\Gamma}_{\boldsymbol{\gamma}}$; $(\mathcal{C}_2)$ $T^{-1}\sum_{t=1}^{T}E\left(\mathbf{V}_j\mathbb{I}_{\left(|\mathbf{Z}^*_{\gamma,j}|>\boldsymbol{\epsilon}\sqrt{T}\right)}|\mathcal{G}_{j-1}\right) = o_p(1)$ for $\boldsymbol{\epsilon} > 0$; and $(\mathcal{C}_3)$ $T^{-1}\sum_{j=1}^{T}E(\mathbf{V}_j|\mathcal{G}_{j-1}) \xrightarrow{p} \sigma^4\boldsymbol{\Gamma}_{\boldsymbol{\gamma}}$, where $\mathbf{V}_j \equiv \mathbf{Z}^*_{\gamma,j}\mathbf{Z}^{*\prime}_{\gamma,j}$.

To show that $\mathcal{C}_1$ holds, first note that $\{Z_j, \mathcal{G}_j\}$ is also a MDS such that $E(Z_j Z_s) = \sigma^4 (T - j)$ for all $j = s$, and 0 otherwise. Hence, $T^{-1} E\left(\mathbf{S}_{\boldsymbol{\gamma},T} \mathbf{S}'_{\boldsymbol{\gamma},T}\right) = \sigma^4 \sum_{j=1}^{T} \left[\boldsymbol{\omega}_j \boldsymbol{\omega}'_j\right] (1 - j/T) \to \sigma^4 \boldsymbol{\Gamma}_{\boldsymbol{\gamma}}$ as T diverges because $j\left(\boldsymbol{\omega}_j \boldsymbol{\omega}'_j\right) = O(1/j)$, so the weighting scheme $\boldsymbol{\omega}_j$ provides squared-summability which in turn ensures (Hannan, 1970) $\lim_{T\to\infty} T^{-1} \sum_{j=1}^{T} \mathbf{S}_{\boldsymbol{\gamma},T} \mathbf{S}'_{\boldsymbol{\gamma},T} \xrightarrow{p} \sigma^4 \boldsymbol{\Gamma}_{\boldsymbol{\gamma}}$. To verify $\mathcal{C}_2$, write $\mathbf{Z}^*_{\gamma,j} = \left(Z^*_{1,j}, ..., Z^*_m\right)'$, and note that the conditional Lindeberg condition must hold on all the entries of the matrix $\mathbf{V}_j$, corresponding to the partial sums of variance terms $\left\{Z^{*2}_{l,j}\right\}$, $l = 1, ..., m$, and the cross-products $\left\{Z^*_{n,j} Z^*_{m,j}\right\}$, $n \neq m$. In order to save space, we show the proof for the diagonal elements $\left\{Z^{*2}_{l,j}\right\}$ and omit the cross-product case, as it follows along the same lines.

For any $1 \leq l \leq k+2$, Cauchy-Schwartz's inequality shows

$$E\left(Z^{*2}_{l,j} \mathbb{I}_{\left(|Z^*_{l,j}| > \epsilon\sqrt{T}\right)} | \mathcal{G}_{j-1}\right) \leq \left[E\left(Z^{*4}_{l,j} | \mathcal{G}_{j-1}\right) E\left(\left(\mathbb{I}_{\left(|Z^*_{l,j}| > \epsilon\sqrt{T}\right)}\right)^2 | \mathcal{G}_{j-1}\right)\right]^{1/2}.$$

Since $E\left(\varepsilon_t^4\right) = \kappa < \infty$, the first expectation on the right-hand side is

$$\begin{aligned} E\left(Z^{*4}_{l,j} | \mathcal{G}_{j-1}\right) &= \omega_j^4(\gamma_l) E\left(Z_j^4 | \mathcal{G}_{j-1}\right) = \omega_j^4(\gamma_l) \left(\kappa \sum_{t=1}^{T-j} \varepsilon_t^4 + 6\sigma^2 \sum_{t=1}^{T-j} \varepsilon_t^2\right) + o_p(T) \\ &= \omega_j^4(\gamma_l)\left(R_{T-j}\right), \text{ say.} \end{aligned}$$

For the second expectation term, the properties of conditional expectations and Chebyshev's inequality yield

$$\begin{aligned} E\left(\left(\mathbb{I}_{\left(|Z^*_{l,j}| > \epsilon\sqrt{T}\right)}\right)^2 | \mathcal{G}_{j-1}\right) &= \Pr\left(|\omega_j(\gamma_l) Z_j| > \epsilon\sqrt{T} | \mathcal{G}_{j-1}\right) \\ &= \Pr\left(|Z_j| > \frac{\epsilon\sqrt{T}}{|\omega_j(\gamma_l)|} | \mathcal{G}_{j-1}\right) \leq \frac{\omega_j^2(\gamma_l)\sigma^2}{\epsilon^2 T} \sum_{t=1}^{T-j} \varepsilon_t^2 . \end{aligned}$$

Hence,

$$\begin{aligned}
T^{-1}\sum_{j=1}^{T}E\left(Z_j^{*2}\mathbb{I}_{\left(|Z_j^*|>\epsilon\sqrt{T}\right)}|\mathcal{G}_{j-1}\right) &\leq \frac{\sigma}{\epsilon\sqrt{T}}\sum_{j=1}^{T}\omega_j^3\left(\gamma_l\right)\left\{\frac{R_{T-j}}{T}\sum_{t=1}^{T-j}\varepsilon_t^2/T\right\}^{1/2}+o_p\left(1\right)\\
&\leq \frac{\sigma}{\epsilon\sqrt{T}}\sum_{j=1}^{T}\omega_j^3\left(\gamma_l\right)\left\{\frac{R_{T}}{T}\sum_{t=1}^{T}\varepsilon_t^2/T\right\}^{1/2}+o_p\left(1\right)\\
&\leq \frac{\sigma^2\sqrt{\kappa^2+6\sigma^2}}{\epsilon\sqrt{T}}\sum_{j=1}^{T}\omega_j^3\left(\gamma_l\right)+o_p\left(1\right)\\
&= o_p\left(1\right)
\end{aligned}$$

because $\lim_{T\to\infty}\sum_{j=1}^{T}\omega_j^p\left(\gamma_l\right)=O\left(1\right)$ for all $p\geq 2$, and $\lim_{T\to\infty}R_T/T=\kappa^2+6\sigma^4$.

Finally, to show condition $\mathcal{C}_3$, note that

$$T^{-1}\sum_{j=1}^{T}E\left(\mathbf{V}_j|\mathcal{G}_{j-1}\right)=\sigma^2T^{-1}\sum_{j=1}^{T}\left[\boldsymbol{\omega}_j\boldsymbol{\omega}_j'\right]\left(\sum_{t=1}^{T-j}\varepsilon_t^2\right)\to\sigma^4\mathbf{\Gamma}_{\boldsymbol{\gamma}}+o\left(1\right)$$

where the last part follows by noting $\sum_{t=1}^{T-j}E\left(\varepsilon_t^2\right)=\sigma^2\left(T-j\right)$, and again, squared-sumability leads to $T^{-1}\sum_{j=1}^{T}E\left(\mathbf{V}_j|\mathcal{G}_{j-1}\right)\xrightarrow{p}\sigma^4\mathbf{\Gamma}_{\boldsymbol{\gamma}}$. Therefore, conditions $\mathcal{C}_1$, $\mathcal{C}_2$, and $\mathcal{C}_3$ jointly imply the required result. This completes the proof.■

Proof of Lemma B.2.

i) First note that $E\left(\mathbf{\Omega}_T\right)=\sigma^2T^{-1}\sum_{t=1}^{T}\sum_{j=1}^{t}\boldsymbol{\omega}_j\boldsymbol{\omega}_j'$. Then direct algebra shows

$$\begin{aligned}
E\left(\mathbf{\Omega}_T\right) &= \sigma^2\sum_{j=1}^{T}\boldsymbol{\omega}_j\boldsymbol{\omega}_j'-\sigma^2T^{-1}\sum_{j=2}^{T}j\left[\boldsymbol{\omega}_j\boldsymbol{\omega}_j'\right]+\sigma^2T^{-1}\sum_{j=2}^{T}\boldsymbol{\omega}_j\boldsymbol{\omega}_j'\\
&= \sigma^2\sum_{j=1}^{T}\boldsymbol{\omega}_j\boldsymbol{\omega}_j'-o\left(1\right)+O\left(T^{-1}\right)\\
&\to \sigma^2\mathbf{\Gamma}_{\boldsymbol{\gamma}}.
\end{aligned}$$

Hence, from Hannan (1970)

$$\lim_{T\to\infty}T^{-1}\sum_{t=2}^{T}\left(\widehat{\boldsymbol{\varepsilon}}_{\boldsymbol{\gamma},t-1}^{*}\widehat{\boldsymbol{\varepsilon}}_{\boldsymbol{\gamma},t-1}^{*\prime}\right)\xrightarrow{p}\sigma^2\mathbf{\Gamma}_{\boldsymbol{\gamma}}$$

$ii)$ $\widehat{\sigma}_e^2 = \sum_{t=2}^{T} \hat{e}_t^2/T$, with the residuals computed as any of the following: $\left\{\varepsilon_t - \hat{\phi}_s \widehat{\varepsilon}^*_{\gamma_s,t-1}\right\}$, $\left\{\varepsilon_t - \hat{\phi} \sum_{s=1}^{m} \widehat{\varepsilon}^*_{\gamma_s,t-1}\right\}$, $\left\{\varepsilon_t - \sum_{s=1}^{m} \hat{\phi}_l \widehat{\varepsilon}^*_{\gamma_s,t-1}\right\}$ under the suitable restriction. Since, under the restriction implied by the null hypothesis the coefficients in these regressions should be zero, we have from the consistency property of the least-squares procedure

$$T^{-1}\sum_{t=2}^{T} \hat{e}_t^2 = T^{-1}\sum_{t=2}^{T} \varepsilon_t^2 + o_p(1) \xrightarrow{p} \sigma^2$$

since $\{\varepsilon_t^2 - \sigma^2\}$ is a stationary and ergodic MDS under Assumption $\mathcal{A}$. This completes the proof.■

Proof of Theorem 3.1.

The proof of the convergence of $\Upsilon^{(n)}$ is immediate in view of Lemma B.1iii) and Lemma B.2. Define $\mathcal{D}_T = \sum_{t=2}^{T} \widehat{\varepsilon}_t \left(\widehat{\boldsymbol{\varepsilon}}^*_{\boldsymbol{\gamma},t-1}\right)/\sqrt{T} = \mathbf{S}_{\boldsymbol{\gamma},T}/\sqrt{T}$ and recall that $\boldsymbol{\Omega}_T = T^{-1}\sum_{t=2}^{T}\left(\widehat{\boldsymbol{\varepsilon}}^*_{\boldsymbol{\gamma},t-1}\widehat{\boldsymbol{\varepsilon}}^{*\prime}_{\boldsymbol{\gamma},t-1}\right)$. Hence, we can write

$$\Upsilon^{(n)} = \left(\frac{1}{\widehat{\sigma}_e^2}\right)\mathcal{D}_T'\boldsymbol{\Omega}_T^{-1}\mathcal{D}_T'.$$

From the invariance principle in Lemma B.1iii), $\mathcal{D}_T \Rightarrow \mathcal{N}\left(0, \sigma^2\boldsymbol{\Gamma}_{\boldsymbol{\gamma}}\right)$, and from the results in Lemma B.2, $\widehat{\sigma}_e^2 \xrightarrow{p} \sigma^2$, and $\boldsymbol{\Omega}_T^{-1} \xrightarrow{p} \boldsymbol{\Gamma}_{\boldsymbol{\gamma}}^{-1}/\sigma^2$. The required convergence then follows from the Continuous Mapping Theorem (CMT) showing that $\Upsilon^{(n)} \Rightarrow \mathbf{Z}'\mathbf{Z}$, where $\mathbf{Z}$ is a n-dimensional standard normal distribution, and hence $\Upsilon^{(n)} \Rightarrow \chi^2_{(n)}$.

The proof for the cases studied in Corollary 3.1 and Corollary 3.2. also obtains easily from the Lemmas B.1 and B.2 and the CMT. Note that

$$\begin{aligned}\Upsilon_{\gamma_s} &= \frac{1}{\widehat{\sigma}_e^2}\left(\sum_{t=2}^{T}\widehat{\varepsilon}_t\widehat{\varepsilon}^*_{\gamma_s,t-1}/\sqrt{T}\right)^2\left(\sum_{t=2}^{T}\widehat{\varepsilon}^{*2}_{\gamma_s,t-1}/T\right)^{-1}\\ &= \frac{1}{\widehat{\sigma}_e^2}\left(A_{\gamma_s,T}\right)^2\ \left(B_{\gamma_s,T}\right)^{-1}, \text{ say.}\end{aligned}$$

Note that $A_{\gamma_s,T} = \mathbf{1}'_{m,s}\mathcal{D}_T$, where $\mathbf{1}_{m,s}$ is a vector in R^m that takes value one in the s-th entry and zero otherwise. From this, it is easy to show that $A_{\gamma_s,T}$ has zero mean and asymptotic variance $\sigma^2\mathbf{1}'_{m,s}\boldsymbol{\Gamma}_{\boldsymbol{\gamma}}\mathbf{1}_{m,s} = \sigma^4\psi\left(\gamma_s\right)$, and from Lemma B.1iii), $A_{\gamma_s,T} \Rightarrow \sigma^2\sqrt{\psi\left(\gamma_s\right)}\mathcal{N}(0,1)$ as $T \to \infty$. Similarly, $B_{\gamma_s,T} = \mathbf{1}'_{m,s}\boldsymbol{\Omega}_T\mathbf{1}_{m,s}$, so from Lemma B.2, $B_{\gamma_s,T} \xrightarrow{p} \sigma^2\psi\left(\gamma_s\right)$, and $\widehat{\sigma}^2_{e,s} \xrightarrow{p} \sigma^2$. Hence, the CMT completes the proof. For the joint restricted test statistic discussed in Corollary 3.2, the proof is totally analogous by noting $\bar{\Upsilon}^{(n)} = \left[\sigma^2 \mathbf{1}_n' \mathbf{\Omega}_T \mathbf{1}_n\right]^{-1} \left[\mathbf{1}_n' \mathcal{D}_T\right]^2 / \widehat{\sigma}_e^2$, so the results in Lemmas B.1 and B.2, together with the CMT, render the required convergence as $T \to \infty$.■

Proof of Theorem 3.2.

Under Assumption $\mathcal{A}^*$, the proofs are similar to Demetrescu *et al.* (2006) [DKH henceforth] by noting that the context studied $\gamma_s = 0$ can be generalized for $\gamma_s \in [0, \pi]$. For instance, under linear errors and the restrictions in DKH, Lemma B.1 in that paper easily generalizes, noting that $\widehat{\varepsilon}^*_{\gamma_s, t-1} = \sum_{j=1}^{t-j} \omega(j) \varepsilon_{t-j} = \sum_{j=0}^{t-j} \varphi_j u_{t-1-j}$ is square summable for any $\gamma_s \in [0, \pi]$.

References

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

[2] Agliakloglou, C. and P. Newbold (1994) Lagrange multiplier tests for fractional difference. Journal of Time Series Analysis 14, 253–262.

[3] Arteche, J. (2002), Semiparametric robust tests on seasonal or cyclical long memory time series, Journal of Time Series Analysis 23, 251–285.

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

[5] Caporale, G.M. and Gil-Alaña, L.A. (2006) Testing for unit and fractional orders of integration in the trend and seasonal components of US monetary aggregates. Economics and Finance Working Papers, Brunel University 06-13.

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

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

[8] Demetrescu, M., V. Kuzin, U. Hassler (2007) Long Memory Testing in the Time Domain. Forthcoming Econometric Theory.

[9] Gil-Alaña, L.A. and P.M. Robinson (2000) Testing of Seasonal Fractional Integration in UK and Japanese Consumption and Income, London School of Economics, Discussion Paper No. EM/00/402.

[10] Giriatis, L., J. Hidalgo and P.M. Robinson (2001) Gaussian estimation of parametric spectral density with unknown pole. Annals of Statistics, 29, 987-1023.

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

[12] Hassler, U. and J. Breitung (2002) Inference on the Cointegration Rank in Fractionally Integrated Processes, Journal of Econometrics, 110(2), 167-185.

[13] Hassler, U. (1994) (Mis)specification of Long Memory in Seasonal Time Series. Journal of Time Series Analysis 15, 19–30.

[14] Hidalgo, J., and P. Soulier (2004) Estimation of the location and the exponent of the spectral singularity of a long memory process. Journal of Time Series Analysis 25, 55-81.

[15] Hidalgo, J. (2005) Semiparametric estimation for stationary processes whose spectra have an unkown pole. Annals of Statistics 33, 1843-1889.

[16] Nielsen, M.∅. (2004) Efficient likelihood inference in nonstationary univariate models. Econometric Theory 20, 116-146.

[17] Ramachandran, R. and P. Beaumont (2001) Robust Estimation of GARMA Model Parameters with Application to Cointegration Among Interest Rates of Industrialized Countries. Computational Economics 17, 179 - 201.

[18] Smallwood, A.D. and S.C. Norrbin (2005) Generalized Long Memory and the Purchasing Power Parity Controversy, Discussion Paper.

[19] Smallwood, A.D. and S.C. Norrbin (2006) Generalized Long Memory Processes, Failure of Cointegration Tests and Exchange Rate Dynamics, Journal of Applied Econometrics 21, 409 - 417.

[20] Tanaka, K. (1999) The Nonstationary Fractional Unit Root. Econometric Theory 15, 549 - 582.

Tables

Table 1: Size and Power when the DGP is a Simple GARMA model

	θ						
Frequency	-0.3	-0.2	-0.1	0	0.1	0.2	0.3
				T=100			
$\frac{\pi}{10}$	.999	.984	.540	.052	.584	.981	.999
$\frac{2\pi}{10}$	.999	.933	.401	.054	.445	.927	.998
$\frac{3\pi}{10}$	.988	.810	.302	.056	.329	.832	.982
$\frac{4\pi}{10}$	.946	.689	.232	.049	.267	.721	.946
$\frac{5\pi}{10}$	.929	.630	.210	.050	.248	.686	.932
$\frac{6\pi}{10}$	.955	.683	.236	.051	.269	.730	.947
$\frac{7\pi}{10}$	.985	.826	.311	.045	.331	.836	.985
$\frac{8\pi}{10}$	.998	.929	.425	.051	.452	.933	.998
$\frac{9\pi}{10}$	.999	.982	.536	.050	.585	.984	.999
				T=250			
$\frac{\pi}{10}$	.999	.999	.924	.043	.921	.999	.999
$\frac{2\pi}{10}$	.999	.999	.818	.057	.814	.999	.999
$\frac{3\pi}{10}$	.999	.997	.653	.050	.686	.995	.999
$\frac{4\pi}{10}$	.999	.979	.516	.052	.563	.980	.999
$\frac{5\pi}{10}$	.999	.971	.468	.051	.545	.968	.999
$\frac{6\pi}{10}$	.999	.980	.520	.051	.571	.978	.999
$\frac{7\pi}{10}$	.999	.998	.664	.045	.682	.994	.999
$\frac{8\pi}{10}$	.999	1.00	.811	.050	.816	.999	.999
$\frac{9\pi}{10}$	.999	.999	.918	.045	.913	.999	.999

Table 2: Size and Power when the DGP is a 2-factor GARMA model ($\theta_1 = 0.15$ and $\theta_2 = \frac{\pi}{2}$) - T=100

Test on θ_1

θ_1 \ θ_2	-0.3	-0.2	-0.1	0	0.1	0.2	0.3
-0.3	0.999	0.999	0.999	0.999	0.999	0.999	0.999
-0.2	0.975	0.981	0.987	0.992	0.999	0.999	0.999
-0.1	0.411	0.452	0.508	0.608	0.707	0.827	0.924
0.0	0.143	0.098	0.073	0.053	0.062	0.103	0.244
0.1	0.868	0.812	0.751	0.633	0.488	0.282	0.123
0.2	0.999	0.998	0.995	0.988	0.975	0.936	0.817
0.3	0.999	0.999	0.999	0.999	0.999	0.999	0.994

2-lags Augmented Test on θ_1

θ_1 \ θ_2	-0.3	-0.2	-0.1	0	0.1	0.2	0.3
-0.3	0.636	0.714	0.815	0.883	0.950	0.978	0.993
-0.2	0.361	0.393	0.460	0.514	0.604	0.698	0.762
-0.1	0.182	0.165	0.159	0.152	0.145	0.156	0.165
0.0	0.088	0.069	0.052	0.047	0.048	0.053	0.068
0.1	0.062	0.059	0.061	0.098	0.158	0.269	0.400
0.2	0.070	0.054	0.076	0.140	0.286	0.489	0.695
0.3	0.094	0.066	0.068	0.125	0.267	0.511	0.756

Test on θ_2

θ_1 \ θ_2	-0.3	-0.2	-0.1	0	0.1	0.2	0.3
-0.3	0.772	0.430	0.114	0.071	0.364	0.774	0.955
-0.2	0.756	0.375	0.088	0.072	0.398	0.799	0.965
-0.1	0.814	0.408	0.107	0.061	0.358	0.778	0.954
0.0	0.923	0.625	0.202	0.046	0.253	0.660	0.929
0.1	0.994	0.912	0.597	0.187	0.113	0.444	0.814
0.2	0.999	0.997	0.953	0.707	0.308	0.213	0.497
0.3	0.999	0.999	0.999	0.976	0.835	0.502	0.318

2-lags Augmented Test on θ_2

θ_1 \ θ_2	-0.3	-0.2	-0.1	0	0.1	0.2	0.3
-0.3	0.584	0.370	0.157	0.074	0.082	0.174	0.313
-0.2	0.646	0.362	0.176	0.073	0.079	0.178	0.311
-0.1	0.638	0.361	0.154	0.057	0.076	0.190	0.344
0.0	0.601	0.312	0.116	0.045	0.097	0.249	0.417
0.1	0.554	0.260	0.083	0.046	0.129	0.322	0.528
0.2	0.539	0.232	0.064	0.043	0.180	0.424	0.661
0.3	0.610	0.284	0.087	0.049	0.191	0.475	0.710

Joint Restricted Test

θ_1 \ θ_2	-0.3	-0.2	-0.1	0	0.1	0.2	0.3
-0.3	1.000	1.000	0.997	0.959	0.741	0.362	0.247
-0.2	0.996	0.992	0.963	0.834	0.512	0.220	0.237
-0.1	0.793	0.731	0.611	0.398	0.179	0.098	0.290
0.0	0.126	0.102	0.082	0.047	0.067	0.205	0.480
0.1	0.631	0.590	0.583	0.574	0.625	0.730	0.853
0.2	0.987	0.985	0.982	0.981	0.982	0.988	0.993
0.3	0.999	0.999	0.999	0.999	0.999	0.999	0.999

Joint Unrestricted Test

θ_1 \ θ_2	-0.3	-0.2	-0.1	0	0.1	0.2	0.3
-0.3	0.999	0.999	0.999	0.999	0.999	0.999	0.999
-0.2	0.994	0.978	0.974	0.977	0.990	0.999	0.999
-0.1	0.892	0.684	0.502	0.487	0.693	0.911	0.985
0.0	0.857	0.510	0.161	0.049	0.205	0.592	0.893
0.1	0.988	0.913	0.741	0.556	0.535	0.718	0.898
0.2	0.999	0.999	0.992	0.980	0.974	0.981	0.991
0.3	0.999	0.999	0.999	0.999	0.999	0.999	0.999

Table 3: Size and Power when the DGP is a 2-factor GARMA model with ARMA errors ($\theta_1 = 0.15$ and $\theta_2 = \frac{\pi}{2}$) - T=100

All tests augmented, Schwertz's rule

Test on θ_1

θ_1 \ θ_2	-0.3	-0.2	-0.1	0	0.1	0.2	0.3
-0.3	0.078	0.105	0.162	0.215	0.285	0.373	0.427
-0.2	0.032	0.046	0.063	0.090	0.145	0.173	0.214
-0.1	0.034	0.031	0.045	0.044	0.061	0.079	0.093
0.0	0.050	0.049	0.046	0.045	0.042	0.046	0.053
0.1	0.077	0.069	0.068	0.061	0.059	0.060	0.057
0.2	0.097	0.095	0.100	0.090	0.092	0.093	0.093
0.3	0.127	0.127	0.125	0.134	0.132	0.134	0.145

Test on θ_2

θ_1 \ θ_2	-0.3	-0.2	-0.1	0	0.1	0.2	0.3
-0.3	0.160	0.070	0.035	0.051	0.096	0.197	0.276
-0.2	0.181	0.088	0.048	0.049	0.086	0.162	0.243
-0.1	0.199	0.098	0.056	0.042	0.065	0.124	0.182
0.0	0.183	0.113	0.060	0.043	0.053	0.087	0.134
0.1	0.150	0.100	0.062	0.047	0.050	0.060	0.085
0.2	0.097	0.071	0.055	0.039	0.043	0.053	0.065
0.3	0.063	0.059	0.043	0.042	0.044	0.046	0.054

Joint Restricted Test

θ_1 \ θ_2	-0.3	-0.2	-0.1	0	0.1	0.2	0.3
-0.3	0.300	0.233	0.148	0.088	0.067	0.099	0.142
-0.2	0.131	0.120	0.089	0.059	0.051	0.072	0.127
-0.1	0.063	0.056	0.055	0.045	0.041	0.057	0.096
0.0	0.047	0.043	0.046	0.043	0.049	0.062	0.080
0.1	0.065	0.059	0.063	0.060	0.061	0.075	0.086
0.2	0.093	0.087	0.092	0.094	0.092	0.104	0.113
0.3	0.126	0.127	0.123	0.136	0.127	0.130	0.139

Joint Unrestricted Test

θ_1 \ θ_2	-0.3	-0.2	-0.1	0	0.1	0.2	0.3
-0.3	0.204	0.142	0.122	0.141	0.202	0.315	0.381
-0.2	0.156	0.097	0.058	0.069	0.115	0.160	0.228
-0.1	0.137	0.075	0.046	0.039	0.058	0.094	0.138
0.0	0.121	0.076	0.046	0.037	0.044	0.063	0.090
0.1	0.113	0.079	0.058	0.053	0.053	0.062	0.075
0.2	0.103	0.077	0.073	0.061	0.068	0.075	0.085
0.3	0.105	0.094	0.085	0.096	0.091	0.100	0.105