Granger Centre
for time series econometrics

Granger Centre Discussion Paper Series

On augmented HEGY tests for seasonal unit roots

by

Tomás del Barrio Castro, Denise R. Osborn
and A. M. Robert Taylor

Granger Centre Discussion Paper No. 11/01

The University of Nottingham

On Augmented HEGY Tests for Seasonal Unit Roots

Tomás del Barrio Castro
Department of Applied Economics
University of the Balearic Islands

Denise R. Osborn
Economics, School of Social Sciences
University of Manchester

A.M. Robert Taylor
School of Economics
University of Nottingham

June 2011

Abstract

The contribution of this paper is twofold. First we extend the large sample results provided for the augmented Dickey-Fuller test by Said and Dickey (1984) and Chang and Park (2002) to the case of the augmented seasonal unit root tests of Hylleberg *et al.* (1990) [HEGY], *inter alia.* Our analysis is performed under the same conditions on the innovations as in Chang and Park (2002), thereby allowing for general linear processes driven by (possibly conditionally heteroskedastic) martingale difference innovations. We show that the limiting null distributions of the t-statistics for unit roots at the zero and Nyquist frequencies and joint F-type statistics are pivotal, while those of the t-statistics at the harmonic seasonal frequencies depend on nuisance parameters which derive from the lag parameters characterising the linear process. Moreover, the rates on the lag truncation required for these results to hold are shown to coincide with the corresponding rates given in Chang and Park (2002); in particular, an $o(T^{1/2})$ rate is shown to be sufficient. The second contribution of the paper is to explore the use of data-dependent lag selection methods in the context of the augmented HEGY tests. Information criteria based methods along with sequential rules, such as those of Ng and Perron (1995) and Beaulieu and Miron (1993), are compared.

Keywords: seasonal unit root; HEGY tests; linear process; autoregressive approximation; data-based lag selection.

JEL code: C22.

1 Introduction

This paper considers testing for *seasonal* unit roots in a univariate time-series process. In the seminal paper in the literature, Hylleberg *et al.* (1990) [HEGY] develop separate regression-based t- and F-tests for unit roots at the zero, Nyquist and annual (harmonic) frequencies in the context of quarterly data. Recently, Smith, Taylor and del Bario Castro (2009) have generalised the HEGY approach to allow for an arbitrary seasonal aspect.

In the non-seasonal context, Said and Dickey (1984) establish that the Augmented Dickey-Fuller [ADF] statistics have (pivotal) Dickey-Fuller limiting null distributions in the presence of ARMA processes of unknown finite order with independently and identically distributed (*iid*) zero mean and constant variance innovations. This result requires that the lag length employed should increase in proportion with the sample size T at rate $o(T^{1/3})$. More recently, Chang and Park (2002) extend these results to allow the shocks to follow a general linear process driven by potentially heteroskedastic martingale difference innovations and show, moreover, that a weaker $o(T^{1/2})$ rate condition is sufficient for the result to hold. To the best of our knowledge, however, analogous results have not been established for the HEGY-type tests. Previous contributions have either assumed that the shocks are serially uncorrelated (e.g. Smith and Taylor, 1998), or follow a finite-order autoregressive (AR) process (e.g. Burridge and Taylor, 2001, Rodrigues and Taylor, 2004a, 2004b and Smith *et al.*, 2009) or a finite-order moving average (MA) process (del Barrio Castro and Osborn, 2011). Therefore, none allow for either the finite-order ARMA assumption of Said and Dickey (1984) or the more general linear process assumptions of Chang and Park (2002). As Taylor (2005, p.34) notes "It has been widely conjectured, but never formally proved, that an approach along the lines of that developed in regard of the ADF test by Said and Dickey (1984) will purge the effects of ARMA behavior in the shocks from the limiting null distribution of the augmented HEGY-type statistics."

It has been known since the seminal work of Box and Jenkins (1976) that seasonally observed time series can display moving average behaviour, in addition to seasonal (autoregressive) unit roots. Indeed, Box and Jenkins (1976) developed the well-known seasonal ARIMA factorisations (a prominent example of which is the so-called *airline model*) as a parsimonious device for modelling dependence in seasonal data. Allowing for both autoregressive and moving average behaviour is, therefore, very important when testing for unit roots in a seasonal context. ARMA behaviour in the shocks can also be a manifestation of neglected periodic autoregressive (PAR) behaviour, something which has been largely overlooked in the context of the HEGY tests. A detailed discussion of the PAR class of processes is given in Ghysels and Osborn (2001, Chapter 6). As an example, the first-order stationary PAR process for a series observed with period S, denoted $PAR(1)_S$, admits a stationary and invertible ARMA representation, which combines a seasonal autoregressive term with an MA$(S-1)$ component.

Motivated by these considerations, this paper makes two key contributions. First we demonstrate that the results given by Said and Dickey (1984) and Chang and Park (2002) for the ADF statistic do extend, at least in part, to the case of the augmented HEGY-type tests. Specifically, and complementing the findings of previous authors for finite order shocks of either AR or MA form, we show that, provided the order of the lag augmentation polynomial increases in proportion with the sample size at a suitable rate, then the t-statistics for unit roots at the zero and Nyquist frequencies, together with all the F-type tests employed in this context, remain pivotal in the presence of general linear processes driven by martingale difference innovations. However, this is not the case for the t-statistics at the harmonic seasonal frequencies, where the asymptotic null distributions depend on nuisance parameters which derive from the coefficients characterising the linear process. Moreover, we show that the rate restriction required on the order of the lag augmentation polynomial coincides with the rate given in Chang and Park (2002). Empirical practice, however, takes matters a stage further and employs data-dependent methods to select the lag augmentation polynomial. In response to this, the second contribution of the paper is to use Monte Carlo methods to explore the performance of a variety of data-dependent lag selection procedures. These include the sequential test procedures employed by Hall (1994) and Ng and Perron (1995), seasonal variants of these as suggested by Rodrigues and Taylor (2004a) and Beaulieu and Miron (1993), and methods based on information criteria including AIC, BIC and a seasonal generalisation we develop for the modified information criteria [$MAIC$, $MBIC$] of Ng and Perron (2001).

The remainder of the paper is organised as follows. In section 2 we outline the seasonal model, define the hypotheses of interest within that model, and briefly review the augmented HEGY-type seasonal unit root tests. The limiting null distributions of the HEGY statistics when the shocks follow a general linear process driven by martingale difference innovations is established in section 3, along with the necessary rate conditions. Our investigation of the relative finite sample performance of quarterly augmented HEGY tests which arise from a variety of data-based methods of lag selection for series driven by both MA and AR shocks is considered in section 4. Section 5 concludes. Proofs of our main results are contained in a mathematical appendix.

In the following '$\xrightarrow{d}$' denotes weak convergence and '$\xrightarrow{p}$' convergence in probability, in each case as the sample size diverges to positive infinity; '$a := b$' ('$a =: b$') indicates that a is defined by b (b is defined by a); $\lfloor \cdot \rfloor$ denotes the integer part of its argument, and I_p denotes the $p \times p$ identity matrix. The Euclidean norm of the $k \times 1$ vector, x, is defined as $\|x\| := (x'x)^{1/2}$, while for the $k \times k$ matrix, A, we also define $\|A\| := \max_x \|Ax\| / \|x\|$. Finally we define $\mathrm{i} := \sqrt{(-1)}$.

2 The Seasonal Unit Root Framework

2.1 The Seasonal Model and Assumptions

Consider the univariate seasonal time-series process $\{x_{St+s}\}$ which satisfies the following data generating process (DGP)

$$\alpha(L)x_{St+s} = u_{St+s}, \; s = 1-S, ..., 0, \;\; t = 1, 2, \ldots, N \tag{2.1a}$$

$$u_{St+s} = \psi(L)\varepsilon_{St+s} \tag{2.1b}$$

where the positive integer S denotes the number of seasons[1], and $\alpha(z) := 1 - \sum_{j=1}^{S} \alpha_j^* z^j$, is an $AR(S)$ polynomial in the conventional lag operator, L. The error process u_{St+s} in (2.1b) is taken to be a linear process with $\psi(z) := 1 + \sum_{j=1}^{\infty} \psi_j z^j$. Precise conditions on this process are given below. The *initial conditions*, $x_{1-S}, ..., x_0$, are taken to be of $o_p(T^{1/2})$. In what follows we define $T := SN$.

Following Chang and Park (2002), we make the following assumptions on the innovation sequence $\{\varepsilon_{St+s}\}$ and on the coefficients of the polynomial $\psi(L)$.

[1] So that, for example, $S = 4$ yields the case of quarterly data, $S = 12$ monthly data, and $S = 1$ non-seasonal data.

Assumption A.1 *Let* $(\varepsilon_{St+s}, \mathcal{F}_{St+s})$ *be a martingale difference sequence, with filtration* $(\mathcal{F}_{St+s})$, *where* $\mathcal{F}_{St+s} \subset \mathcal{F}_{St+s+1}$, *for all* s, t, *and such that: (a)* $E\left[\varepsilon_{St+s}^2\right] = \sigma^2$, *(b)* $1/N \sum_{t=1}^{N} \varepsilon_{St+s}^2 \xrightarrow{p} \sigma^2$ *for each* $s = 1-S, ..., 0$, *and (c)* $E\left|\varepsilon_{St+s}\right|^r < K$ *with* $r \geq 4$, *where* K *is some constant depending only upon* r.

Assumption A.2 *Let the polynomial* $\psi(z)$ *be such that: (a)* $\psi(z) \neq 0$ *for all* $|z| \leq 1$, *and (b)* $\sum_{j=1}^{\infty} |j|^{\tau} \left|\psi_j\right| < \infty$ *for some* $\tau \geq 1$.

For some of the results in this paper, it will be necessary to substitute Assumption A.1 by the following somewhat stronger assumption.

Assumption A.1' *Let* $(\varepsilon_{St+s}, \mathcal{F}_{St+s})$ *be a martingale difference sequence, with filtration* $(\mathcal{F}_{St+s})$, *where* $\mathcal{F}_{St+s} \subset \mathcal{F}_{St+s+1}$, *for all* s, t, *such that: (a)* $E\left(\varepsilon_{St+s}^2 | \mathcal{F}_{St+s-1}\right) = \sigma^2$, *and (b)* $E\left|\varepsilon_{St+s}\right|^r < K$ *with* $r \geq 4$, *where* K *is some constant depending only upon* r.

Remark 1: Assumptions A.1, A.2 and A.1' correspond to Assumptions 1, 2 and 1', respectively, in Chang and Park (2002) - albeit condition (b) of Assumption A.1 is slightly stronger than condition (b) of Assumption 1 in Chang and Park (2002), reflecting the seasonal aspect of the data - and we therefore refer the reader to Chang and Park (2002, pp.433-434) for a detailed discussion concerning these. However, notice, in particular, that a special case of Assumption A.2 is where u_{St+s} in (2.1b) admits the causal and invertible ARMA(p, q) representation, $\phi(L)u_{St+s} = \theta(L)\varepsilon_{St+s}$, such that all the roots of $\phi(z) := 1 - \sum_{i=1}^{p} \phi_p z^i$ and $\theta(z) := 1 - \sum_{i=1}^{q} \theta_i z^i$ lie strictly outside the unit circle. Assumption A.1 is also weaker than those previously made about the innovation process, ε_{St+s}, in the regression-based seasonal unit root literature, where either Assumption A.1' or the even stronger assumption that ε_{St+s} is $IID(0, \sigma^2)$ with finite fourth moment has been adopted.

2.2 The Seasonal Unit Root Hypotheses

Our focus is on tests for seasonal unit roots in $\alpha(L)$ of (2.1a); that is, the null hypothesis of interest is

$$H_0 : \alpha(L) = 1 - L^S =: \Delta_S. \tag{2.2}$$

Under H_0 of (2.2), the DGP (2.1) of $\{x_{St+s}\}$ is a seasonally integrated process. We may factorise the AR(S) polynomial $\alpha(L)$ as $\alpha(L) = \prod_{j=0}^{\lfloor S/2 \rfloor} \omega_j(L)$, where $\omega_0(L) := (1 - \alpha_0 L)$ associates the parameter α_0 with the zero frequency $\omega_0 := 0$, $\omega_j(L) := [1 - 2(\alpha_j \cos\omega_j - \beta_j \sin\omega_j)L + (\alpha_j^2 + \beta_j^2)L^2]$ corresponds to the conjugate (harmonic) seasonal frequencies $(\omega_j, 2\pi - \omega_j)$, $\omega_j := 2\pi j/S$, with associated parameters α_j and β_j, $j = 1, ..., S^*$, where $S^* := \lfloor (S-1)/2 \rfloor$, and, for S even, $\omega_{S/2}(L) := (1 + \alpha_{S/2} L)$, associates the parameter $\alpha_{S/2}$ with the Nyquist frequency[2] $\omega_{S/2} := \pi$. Consequently H_0 of (2.2) may be commensurately partitioned as $H_0 = \cap_{j=0}^{\lfloor S/2 \rfloor} H_{0,j}$, where

$$H_{0,i} : \alpha_i = 1, \;\; i = 0, \; S/2, \;\; \text{and} \;\; H_{0,j} : \alpha_j = 1, \; \beta_j = 0, \;\; j = 1, ..., S^*. \tag{2.3}$$

[2]As a point of notation, throughout the paper where reference is made to the Nyquist frequency this is understood only to apply where S is even. Where S is odd, elements and discussion pertaining to the Nyquist frequency should simply be deleted.

The hypothesis $H_{0,0}$ corresponds to a unit root at the zero frequency while $H_{0,S/2}$ yields a unit root at the Nyquist frequency, $\omega_{S/2} = \pi$. A pair of complex conjugate unit roots at the jth harmonic seasonal frequencies is obtained under $H_{0,j}$, $j = 1, ..., S^*$.

The alternative hypothesis H_1 is of stationarity at one or more of the zero or seasonal frequencies; that is, $H_1 = \cup_{j=0}^{\lfloor S/2 \rfloor} H_{1,j}$, where

$$H_{1,i} : \alpha_i < 1, \quad i = 0, S/2, \quad \text{and} \quad H_{1,j} : \alpha_j^2 + \beta_j^2 < 1, \quad j = 1, ..., S^*. \tag{2.4}$$

Consequently, the maintained hypothesis $H_0 \cup H_1$ excludes all unit roots, except for a possible single unit root at each of the zero and Nyquist frequencies and a single pair of complex conjugate unit roots at each of the harmonic seasonal frequencies. Explosive roots in $\alpha(L)$ are also excluded.

2.3 The Augmented HEGY Tests

In order to develop regression-based seasonal unit root tests, note first that under Assumption A.2 $\psi(z)$ is invertible, and let the (unique) inverse of $\psi(z)$ be denoted $d(z) := 1 - \sum_{j=1}^{\infty} d_j z^j$. Then, it follows, using the Proposition in HEGY (1990, pp.221-222) that expanding $\alpha(z)d(z)$ around the zero and seasonal frequency unit roots, $\exp(\pm \mathrm{i} 2\pi j/S)$, $j = 0, ..., \lfloor S/2 \rfloor$, that the hypotheses in (2.3) may be re-stated as $H_{0,0} : \pi_0 = 0$, $H_{0,S/2} : \pi_{S/2} = 0$, and $H_{0,j} : \pi_j = \pi_j^* = 0$, $j = 1, ..., S^*$, in the model

$$d^*(L)\Delta_S x_{St+s} = \pi_0 x_{0,St+s} + \pi_{S/2} x_{S/2,St+s} + \sum_{j=1}^{S^*} \left(\pi_j x_{j,St+s} + \pi_j^* x_{j,St+s}^* \right) + \varepsilon_{St+s} \tag{2.5}$$

omitting the term $\pi_{S/2} x_{S/2,St+s}$ where S is odd, where $d^*(z) := 1 - \sum_{j=1}^{\infty} d_j^* z^j$ is a causal AR polynomial, and where

$$x_{0,St+s} := \sum_{j=0}^{S-1} x_{St+s-j-1}, \quad x_{S/2,St+s} := \sum_{j=0}^{S-1} \cos[(j+1)\pi] x_{St+s-j-1}, \tag{2.6a}$$

$$x_{i,St+s} := \sum_{j=0}^{S-1} \cos[(j+1)\omega_i] x_{St+s-j-1}, \quad x_{i,St+s}^* := -\sum_{j=0}^{S-1} \sin[(j+1)\omega_i] x_{St+s-j-1}, \quad i = 1, ..., S^*. \tag{2.6b}$$

As in Chang and Park (2002, p.434), we can approximate u_{St+s} from (2.1b) in rth mean by the finite-order AR process

$$u_{St+s} = d_1 u_{St+s-1} + \cdots + d_k u_{St+s-k} + e_{St+s}^k \tag{2.7}$$

with

$$e_{St+s}^k = \varepsilon_{St+s} + \sum_{j=k+1}^{\infty} d_j u_{St+s-j}. \tag{2.8}$$

Using the fact that under H_0 of (2.2), $\Delta_S x_{St+s} = u_{St+s}$ and $d^*(L) = d(L)$, substituting (2.7) into (2.5) yields the auxiliary regression equation

$$\Delta_S x_{St+s} = \pi_0 x_{0,St+s} + \pi_{S/2} x_{S/2,St+s} + \sum_{j=1}^{S^*} \left(\pi_j x_{j,St+s} + \pi_j^* x_{j,St+s}^* \right) + \sum_{j=1}^{k} d_j \Delta_S x_{St+s-j} + e_{St+s}^k \tag{2.9}$$

again omitting the term $\pi_{S/2}x_{S/2,St+s}$ where S is odd, so that (2.9) may be estimated by OLS over observations $St+s=k+1,...,T$. Consequently, as discussed for $S=4$ in HEGY and for general S in Smith *et al.* (2009), tests for the presence (or otherwise) of a unit root at the zero and Nyquist frequencies may be obtained using conventional lower tailed regression t-tests, denoted t_0 and $t_{S/2}$, for the exclusion of $x_{0,St+s}$ and $x_{S/2,St+s}$, respectively, from (2.9). Similarly, the hypothesis of a pair of complex unit roots at the jth harmonic seasonal frequency may be tested by the lower-tailed t_j and two-tailed t_j^* regression t-tests for the exclusion of $x_{j,St+s}$ and $x_{j,St+s}^*$, respectively, or by the (upper-tailed) regression F-test, denoted F_j, for the exclusion of both $x_{j,St+s}$ and $x_{j,St+s}^*$ from (2.9), $j=1,...,S^*$. Ghysels *et al.* (1994) for $S=4$ and Smith *et al.* (1990), again for general S, also consider the joint frequency (upper-tail) regression F-tests from (2.9), namely $F_{1...\lfloor S/2\rfloor}$ for the exclusion of $x_{S/2,St+s}$, together with $x_{j,St+s}$ and $x_{j,St+s}^*$, $j=1,...,S^*$, and $F_{0...\lfloor S/2\rfloor}$ for the exclusion of $x_{0,St+s}$, $x_{S/2,St+s}$, and $x_{j,St+s}$ and $x_{j,St+s}^*$, $j=1,...,S^*$. The former tests the null hypothesis of unit roots at all seasonal frequencies, while the latter tests the overall null, H_0.

As discussed in Chang and Park (2002, p.434), under Assumption A.2 it holds that $\sum_{j=1}^{\infty}|j|^{\tau}|d_j|<\infty$, and consequently $\sum_{j=k+1}^{\infty}|d_j|=o\left(k^{-\tau}\right)$. Hence, as in Chang and Park (2002), the existence of the rth moment of u_{St+s}, which is implied by Assumptions A.1 and A.2, yields that

$$E\left|e_{St+s}^{k}-\varepsilon_{St+s}\right|^{r}\leq E\left|u_{St+s}\right|^{r}\left(\sum_{j=k+1}^{\infty}|d_j|\right)^{r}=o\left(k^{-r\tau}\right).$$

Consequently, the approximation error in (2.7), and hence in (2.9), becomes small as k gets large. However, as in Chang and Park (2002), an assumption is still required concerning the rate permitted on the lag truncation parameter, k, as the sample size increases. Depending on the context, three possible assumptions can be made, as follows:

Assumption A.3 *Let* $k\to\infty$ *and* $k=o\left(T^{1/2}\right)$ *as* $T\to\infty$.

Assumption A.3' *Let* $k\to\infty$ *and* $k=o\left([T/\log T]^{1/2}\right)$ *as* $T\to\infty$.

Assumption A.3" *Let* $k\to\infty$ *and* $k=o\left(T^{1/3}\right)$ *as* $T\to\infty$.

Remark 2: As noted in Chang and Park (2002), Assumptions A.2 and A.3 are considerably weaker that the corresponding assumptions used in Said and Dickey (1984) which, in particular, rule out the possibility of a logarithmic rate on k. The condition that $k=o\left(T^{1/2}\right)$ imposed by Assumption A.3 is, however, not sufficient to guarantee the consistency of the estimators of the coefficients, d_j, $j=1,...,k$, on the lagged dependent variables in (2.9). As we shall show in the next section, this can be achieved by imposing the slower rates of either Assumptions A.3' or Assumption A.3". The rate imposed by the former guarantees consistency under homogeneous martingale difference innovations, as in Assumption A.1', while the rate imposed by the latter is necessary for possibly heterogenous martingale difference innovations, as allowed under Assumption A.1. These slower rates will therefore be required for sequential lag selection methods, such as those discussed in section 4.

3 Asymptotic Results

Under H_0 of (2.2), $\{x_{St+s}\}$ of (2.1) admits the so-called vector of seasons representation

$$X_t = X_{t-1} + U_t, \quad t = 1, 2, ..., N, \tag{3.1}$$

where $X_t := [x_{St-(S-1)}, x_{St-(S-2)}, ..., x_{St}]'$, $t = 0, ..., N$, and $U_t := [u_{St-(S-1)}, u_{St-(S-2)}, ..., u_{St}]'$, $t = 1, ..., N$. As shown in Burridge and Taylor (2001), the error process, U_t satisfies the vector $MA(\infty)$ representation

$$U_t = \sum_{j=0}^{\infty} \mathbf{\Psi}_j E_{t-j} \tag{3.2}$$

where $E_t := [\varepsilon_{St-(S-1)}, \varepsilon_{St-(S-2)}, ..., \varepsilon_{St}]'$ and the $S \times S$ matrices:

$$\mathbf{\Psi}_0 := \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ \psi_1 & 1 & 0 & 0 & \cdots & 0 \\ \psi_2 & \psi_1 & 1 & 0 & \cdots & 0 \\ \psi_3 & \psi_2 & \psi_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \psi_{S-1} & \psi_{S-2} & \psi_{S-3} & \psi_{S-4} & \cdots & 1 \end{bmatrix}$$

and

$$\mathbf{\Psi}_j := \begin{bmatrix} \psi_{jS} & \psi_{jS-1} & \psi_{jS-2} & \psi_{jS-3} & \cdots & \psi_{jS-(S-1)} \\ \psi_{jS+1} & \psi_{jS} & \psi_{jS-1} & \psi_{jS-2} & \cdots & \psi_{jS-(S-2)} \\ \psi_{jS+2} & \psi_{jS+1} & \psi_{jS} & \psi_{jS-1} & \cdots & \psi_{jS-(S-3)} \\ \psi_{jS+3} & \psi_{jS+2} & \psi_{jS+1} & \psi_{jS} & \cdots & \psi_{jS-(S-4)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \psi_{jS+S-1} & \psi_{jS+S-2} & \psi_{jS+S-3} & \psi_{jS+S-4} & \cdots & \psi_{jS} \end{bmatrix}, \quad j = 1, 2, \ldots$$

In Lemma 1 we now provide a multivariate invariance principle for X_t of (3.1)-(3.2). This provides the basic building block for the asymptotic results given in this paper. In Lemma 2 we then establish the large sample properties of the focal variables $x_{j,St+s}$, $j = 0, \ldots, \lfloor S/2 \rfloor$, and $x^*_{i,St+s}$, $i = 1, \ldots, S^*$, from (2.9).

Lemma 1 *Let X_t be generated by (3.1)-(3.2). Then under Assumptions A.1 and A.2,*

$$N^{-1/2} X_{\lfloor rN \rfloor} \xrightarrow{d} \sigma\, \mathbf{\Psi}(1)\, \mathbf{W}(r) := \mathbf{B}(r), \quad r \in [0, 1] \tag{3.3}$$

where $\mathbf{W}(r)$ is a $S\times 1$ standard Brownian motion process and $\mathbf{\Psi}(1) := \sum_{j=0}^{\infty} \mathbf{\Psi}_j$. Notice, therefore, that $\mathbf{B}(r)$ is a $S\times 1$ vector Brownian motion process with variance matrix $\mathbf{\Omega} := \sigma^2\, \mathbf{\Psi}(1)\, \mathbf{\Psi}(1)'$. Moreover, the right member of the convergence result in (3.3) can also be written as $\frac{\sigma}{S}\left[\psi(1)C_0 + \psi(-1)C_{S/2}\right.$ $\left. +2\sum_{i=1}^{S^*} (b_i C_i \;\; + a_i C_i^*)\right] \mathbf{W}(r)$, where $C_0 := \mathsf{Circ}[1, 1, 1, \ldots, 1]$, $C_{S/2} = \mathsf{Circ}[1, -1, 1, \ldots, -1]$, and, for $\omega_i = 2\pi i/S$, $C_i = \mathsf{Circ}\left[\cos(0), \cos(\omega_i), \cos(2\omega_i), \ldots, \cos((S-1)\omega_i)\right]$ and $C_i^* = \mathsf{Circ}\left[\sin(0),\right.$ $\left.\sin((S-1)\omega_i), \sin((S-2)\omega_i), \ldots, \sin(\omega_i)\right]$, $i = 1, \ldots, S^*$, are $S \times S$ circulant matrices, and where

$a_i := \mathcal{I}m(\psi[\exp(\mathrm{i}\omega_i)])$ *and* $b_i := \mathcal{R}e(\psi[\exp(\mathrm{i}\omega_i)])$, $i = 1, ..., S^*$, $\mathcal{R}e(\cdot)$ *and* $\mathcal{I}m(\cdot)$ *denoting the real and imaginary parts of their arguments, respectively.*

Remark 3: Of the circulant matrices appearing in Lemma 1, both C_0 and $C_{S/2}$ have rank one, while C_j and C_j^*, $j = 1, \ldots, S^*$, are all of rank two. For further details on circulant matrices see, for example, Davis (1979), Osborn and Rodrigues (2002), and Smith *et al.* (2009).

Lemma 2 *Let the conditions of Lemma 1 hold. Then for* $X_{j,t} := [x_{j,St-(S-1)}, x_{j,St-(S-2)}, \ldots, x_{j,St}]'$, $j = 0, \ldots, \lfloor S/2 \rfloor$, *and* $X_{i,t}^* := [x_{i,St-(S-1)}^*, x_{i,St-(S-1)}^*, \ldots, x_{i,St}^*]'$, $i = 1, \ldots, S^*$, *we have that*

$$N^{-1/2} X_{0,\lfloor rN \rfloor} \xrightarrow{d} \sigma\, \psi(1)\, C_0 \mathbf{W}(r) \tag{3.4}$$

$$N^{-1/2} X_{S/2,\lfloor rN \rfloor} \xrightarrow{d} \sigma\, \psi(-1)\, C_{S/2} \mathbf{W}(r) \tag{3.5}$$

$$N^{-1/2} X_{i,\lfloor rN \rfloor} \xrightarrow{d} \sigma\left(b_i C_i + a_i C_i^*\right) \mathbf{W}(r), \quad i = 1, ..., S^* \tag{3.6}$$

$$N^{-1/2} X_{i,\lfloor rN \rfloor}^* \xrightarrow{d} \sigma\left(b_i C_i^* - a_i C_i\right) \mathbf{W}(r), \quad i = 1, ..., S^* \tag{3.7}$$

where the vector standard Brownian motion, $\mathbf{W}(r)$, *the constants* a_i *and* b_i, $i = 1, ..., S^*$, *and the circulant matrices,* C_j, $j = 0, \ldots, \lfloor S/2 \rfloor$, *and* C_i^*, $i = 1, \ldots, S^*$, *are as defined in Lemma 1.*

Remark 4: It can be seen from the results in Lemma 2 that the right members of (3.4)-(3.7) are formed from linear combinations of the S independent standard Brownian motions which comprise $\mathbf{W}(r)$. Recalling that C_0 and $C_{S/2}$ both have rank one, whereas C_j and C_j^* for $j = 1, \ldots, S^*$, all have rank two, it is seen that each element of $C_i \mathbf{W}(r)$, $i = 0, S/2$, and of $C_j \mathbf{W}(r)$, $C_j^* \mathbf{W}(r)$ $j = 1, \ldots, S^*$, is, after rescaling, a function of a scalar standard Brownian motion and of two standard Brownian motions, respectively. Moreover, since the products $C_0 C_{S/2}$ and $C_i C_j$ and $C_i C_j^*$, $i = 0, S/2$, $j = 1, ..., S^*$, are all zero matrices, it is seen that these Brownian motions arising from the linear combinations which feature in the right members of (3.4) and (3.5) are independent of one another and of those which arise in the right members of (3.6) and (3.7). Moreover, by virtue of the fact that the products $C_i C_j^*$, $C_i C_j$ and $C_i^* C_j^*$, $i, j = 1, ..., S^*$, $i \neq j$, are also all zero matrices, the pairs of Brownian motions featuring in (3.6) and (3.7) are also seen to be independent across $i = 1, ..., S^*$.

Drawing on the results in Lemmas 1 and 2 we may now state our main result which details the large sample behaviour of the unit root statistics from (2.9) under the general linear process assumptions adopted in this paper.

Proposition 1 *Let the conditions of Lemma 1 hold. Moreover, let Assumption A.3 hold in the auxiliary HEGY regression (2.9). Then the* t_0, $t_{S/2}$ *(S even),* t_j *and* t_j^*, $j = 1, \ldots, S^*$, *statistics from (2.9) are such that:*

$$t_i \xrightarrow{d} \frac{\int_0^1 B_i dB_i}{\sqrt{\int_0^1 B_i^2 dr}} =: \eta_i, \quad i = 0, S/2 \tag{3.8}$$

$$t_j \xrightarrow{d} \frac{a_j\left[\int_0^1 B_j^* dB_j - \int_0^1 B_j dB_j^*\right] + b_j\left[\int_0^1 B_j dB_j + \int_0^1 B_j^* dB_j^*\right]}{\sqrt{(a_j^2 + b_j^2)\left[\int_0^1 B_j^2 dr + \int_0^1 B_j^{*\,2} dr\right]}}, \quad j = 1, ..., S^* \tag{3.9}$$

$$t_j^* \xrightarrow{d} \frac{a_j\left[\int_0^1 B_j^* dB_j^* + \int_0^1 B_j dB_j\right] + b_j\left[\int_0^1 B_j dB_j^* - \int_0^1 B_j^* dB_j\right]}{\sqrt{(a_j^2+b_j^2)\left[\int_0^1 B_j^2 dr + \int_0^1 B_j^{*2} dr\right]}}, \quad j = 1, ..., S^* \tag{3.10}$$

where B_0, $B_{S/2}$, B_j^* and B_j, $j = 1, \ldots, S^*$, are independent standard (scalar) Brownian motions, and the constants, a_j and b_j, $j = 1, ..., S^*$, are as defined in Lemma 1.

Remark 5: For the quarterly case, $S = 4$, the representations given in (3.8)-(3.10) coincide with the corresponding representations given in, for example, Theorem 2.1 of Burridge and Taylor (2001). Burridge and Taylor (2001) derive their results under conditions on the innovation process ε_{St+s} which are analogous to our Assumption A.1' coupled with the much stronger assumption than our Assumption A.2 that the inverse of $\psi(z)$ is finite-ordered; that is, they assume that u_{St+s} in (2.1b) follows an $AR(p)$ process with p finite. Correspondingly, their results require that the lag truncation in (2.9) is a fixed number (i.e., not a function of the sample size) no smaller than p. Other authors have made the same or stronger assumptions than those made in Burridge and Taylor (2001). We have therefore demonstrated that the limiting distributions obtained by previous authors can be maintained under much weaker assumptions on both the serial dependence in the shocks, u_{St+s}, and on the moments of the innovation process, ε_{St+s}, than in this previous literature. Moreover, our results highlight the fact that the conjecture made by a number of these authors that the lag length would need to increase at rate $o(T^{1/3})$ when MA behaviour is permitted in u_{St+s} to obtain this result is in fact more stringent than is necessary.

Remark 6: In the light of Remark 5 it follows immediately, as in Burridge and Taylor (2001), that the F-type statistics, F_j, $j = 1, ..., S^*$, $F_{1...\lfloor S/2 \rfloor}$ and $F_{0...\lfloor S/2 \rfloor}$ from (2.9) have the following limiting null distributions:

$$F_j \xrightarrow{d} \frac{\left[\int_0^1 B_j dB_j + \int_0^1 B_j^* dB_j^*\right]^2 + \left[\int_0^1 B_j dB_j^* - \int_0^1 B_j^* dB_j\right]^2}{2\left(\int_0^1 B_j^2 dr + \int_0^1 B_j^{*2} dr\right)} := \eta_j, \quad j = 1, ..., S^* \tag{3.11}$$

$$F_{1...\lfloor S/2 \rfloor} \xrightarrow{d} \frac{1}{S-1}\left(\eta_2^2 + 2\sum_{j=1}^{S^*} \eta_j\right), \quad F_{0...\lfloor S/2 \rfloor} \xrightarrow{d} \frac{1}{S}\left(\eta_0^2 + \eta_2^2 + 2\sum_{j=1}^{S^*} \eta_j\right) \tag{3.12}$$

omitting η_2^2 from both expressions in (3.12) when S is odd. For $S = 4$, the representations given in (3.8), (3.11) and (3.12) coincide with those given for the case where u_{St+s} is serially uncorrelated in Smith and Taylor (1998, pp.279-280). The limiting null distributions of the t_0, $t_{S/2}$, F_k, $k = 1, ..., S^*$, $F_{1...\lfloor S/2 \rfloor}$ and $F_{0...\lfloor S/2 \rfloor}$ statistics from (2.9) are therefore invariant to the serial correlation nuisance parameters $\{\psi_j\}_{j=1}^{\infty}$ which characterise the serial dependence in u_{St+s}. Previously tabulated asymptotic critical values for the tests based on these statistics may therefore still be used; for example, (3.8) is the standard Dickey–Fuller distribution tabulated in Fuller (1996, Table 10.A.2, p.642).

Remark 7: Regardless of the serial dependence in u_{St+s}, it is seen from the results in Proposition 1 and Remark 6 that the harmonic frequency statistics t_j, t_j^* and F_j are asymptotically independent across

$j = 1, ..., S^*$ and are asymptotically independent of the zero and Nyquist frequency statistics, t_0 and $t_{S/2}$, respectively, under H_0 of (2.2) by virtue of the mutual independence of B_0, B_j, B_j^*, $j = 1, ..., S^*$, and $B_{S/2}$. Moreover, t_0 is asymptotically independent of $t_{S/2}$ and $F_{1...\lfloor S/2 \rfloor}$.

Remark 8: The proof of Proposition 1 includes the result that under Assumptions A.1, A.2 and A.3 the usual OLS residual variance estimator from (2.9), $\hat{\sigma}^2$ say, has the property that $\hat{\sigma}^2 \xrightarrow{p} \sigma^2$. This result is, of course, required for using information criteria based methods to determine the lag length k in (2.9). Such methods will be explored subsequently in section 4.

Remark 9: Under Assumptions A.1, A.2 and A.3" it can be shown, following the same lines as in the proof of Lemma 3.4 in Chang and Park (2002), that the OLS estimator of $\Phi := [d_1, \ldots, d_k]'$ from (2.9) satisfies $\|\hat{\Phi} - \Phi\| = o_p\left(k^{-1/2}\right)$, for N large. Moreover, under Assumptions A.1', A.2 and A.3' it can also be shown that Lemma 3.5 in Chang and Park (2002) carries over to the seasonal case and, hence, that

$$\hat{\Phi} = \Phi + O_p\left(k\left[\frac{\log T}{T}\right]^{1/2}\right) + o_p\left(k^{-\tau}\right)$$

where τ is defined in Assumption A.2. These results show under which assumptions consistency can be established for the estimators of $d_1, \ldots, d_k$, as required to make use of sequential algorithms, such as those of Ng and Perron (1995) and Beaulieu and Miron (1993), to determine the lag length in (2.9). As discussed by Chang and Park (2002, p.439) in the context of the Dickey-Fuller test, the weak condition of Assumption A.3 that $k = o\left(T^{1/2}\right)$ is not sufficient when the augmentation lag length for the HEGY test is selected using a data-based procedure, since this does not deliver consistency for the estimators of $d_1, \ldots, d_k$.

Remark 10: Thus far we have considered the case where the process $\{x_{St+s}\}$ admits no deterministic component. It is straightforward to extend the foregoing results to the case where the series contains deterministic elements. To that end, consider the following generalisation of (2.1):

$$x_{St+s} = \mu_{St+s} + y_{St+s}, \; s = 1 - S, ..., 0, \;\; t = 0, 1, \ldots, N, \tag{3.13a}$$

$$\alpha(L) y_{St+s} = u_{St+s}, \; s = 1 - S, ..., 0, \;\; t = 1, 2, \ldots, N, \tag{3.13b}$$

$$u_{St+s} = \psi(L) \varepsilon_{St+s}. \tag{3.13c}$$

In (3.13a), $\mu_{St+s} := \gamma' Z_{St+s}$ where Z_{St+s} are purely deterministic. The right member of the auxiliary regression in (2.9) must now be correspondingly augmented by the addition of the deterministic component $\mu^*_{St+s} := \gamma^* Z_{St+s}$, where μ_{St+s} and μ^*_{St+s} are linear in the mapping $\gamma \mapsto \gamma^*$. Smith *et al.* (2009) present a typology of six cases of interest for μ_{St+s}, namely: no deterministic component (as considered above); non-seasonal intercept; non-seasonal intercept and non-seasonal trend; seasonal intercepts; seasonal intercepts and non-seasonal trend, and seasonal intercepts and seasonal trends. It is important to notice, as shown in Smith *et al.* (2009), that the inclusion of seasonal intercepts in (2.9) renders the resulting unit root tests similar with respect to the initial conditions, $y_{1-S}, ..., y_0$. Where such deterministic components are included in (2.9), the results given in this section still hold provided the standard Brownian motions, $B_0, B_1, B_1^*, ..., B_{S^*}, B_{S^*}^*$ and (where S is even) $B_{S/2}$, are re-defined as appropriate to the deterministic scenario of interest; cf. Sections 4.1-4.5 of Smith and Taylor (1998) and Smith and Taylor (1999). As an example, if seasonal intercepts are included in (2.9) then the standard Brownian motions above are all replaced by their demeaned analogues, so that (for instance) B_0 is replaced by the process $B_0 - \int_0^1 B_0(s)ds$.

4 Finite Sample Comparison of Lag Selection Methods

Having established the validity of (2.9) under general conditions for the true disturbance process u_{St+s}, this section reports results of a Monte Carlo investigation of the finite sample performance of the quarterly ($S = 4$) HEGY tests when various data-based methods are used to select the lag augmentation polynomial in the test regression. Although results are available for zero frequency unit root tests (see, in particular, Hall, 1994, and Ng and Perron, 1995, 2001), this is not the case in relation to the HEGY tests. Indeed, previous studies that compare the performance of different lag specifications for seasonal unit root tests (including Ghysels *et al.*, 1994, Taylor, 1997, del Barrio Castro and Osborn, 2011) take these lags as given, and hence do not compare the types of data-based procedures often employed by practitioners. After first outlining our methodology, size and (size-adjusted) power results are discussed in the following two subsections.

4.1 Methodology

We generate data according to the model

$$x_{4t+s} = (1 - c/N)x_{4(t-1)+s} + u_{4t+s}, \quad s = -3, ..., 0, \quad t = 1, ..., N \tag{4.1}$$

for both $c = 0$ (size) and $c = 10$ (local power), using $N = 50$, 100. As discussed in Rodrigues and Taylor (2004b), this process is stationary at both the zero and each seasonal frequency when $c > 0$. The initial conditions of the process are set to zero. In addition to white noise innovations, we consider serial correlation in u_{4t+s} of both MA and AR forms, with these being special cases of

$$u_{4t+s} = \left(1 - \theta_1 L - \theta_2 L^2\right)\left(1 - \Theta L^4\right)\varepsilon_{4t+s} \tag{4.2}$$

and

$$(1 - \phi L)\left(1 - \Phi L^4\right) u_{4t+s} = \varepsilon_{4t+s} \tag{4.3}$$

respectively, where ε_{4t+s} is a martingale difference sequence. Our Monte Carlo investigation examined the following cases for (4.2): (i) $\theta_1 = \pm 0.8$, with $\theta_2 = \Theta = 0$; (ii) $\theta_2 = \pm 0.64$, with $\theta_1 = \Theta = 0$; (iii) $\Theta = \pm 0.5, \pm 0.7$ with $\theta_1 = \theta_2 = 0$. For (4.3), the parameterisations considered are: (iv) $\Phi = \pm 0.5$ with $\phi = 0$; (v) $(\phi, \Phi) = (0.8, 0.5), (-0.8, -0.5)$.

Results are reported for the t_0, t_2, F_1, F_{12} and F_{012} test statistics obtained from the quarterly HEGY regression with seasonal intercepts, viz.,

$$\Delta_4 x_{4t+s} = \sum_{s=-3}^{0} \gamma_s^* + \pi_0 x_{0,4t+s} + \pi_2 x_{2,4t+s} + \pi_1 x_{1,4t+s} + \pi_1^* x_{1,4t+s}^* + \sum_{j=1}^{k} d_j \Delta_4 x_{4t+s-j} + e_{4t+s}^k \tag{4.4}$$

where γ_s^* is the intercept for quarter s and e_{4t+s}^k is a disturbance, presumed by the researcher to be IID, when the regression is augmented with k lags of the dependent variable.

The terms of the autoregressive lag augmentation of (4.4) are selected by a variety of data-based methods. First we consider the standard information criteria, AIC and BIC. Second, following Ng and Perron (2001), modified AIC and BIC [$MAIC$ and $MBIC$, respectively] are examined

Under the zero frequency unit root null hypothesis of the augmented Dickey-Fuller [ADF] test, Ng and Perron (2001) extend the Kulback distance measure embedded in information criteria to incorporate the distance from the unit root null; they argue that this additional term is particularly important to account for the presence of a negatively autocorrelated MA process under the null. This leads them to the class of modified information criteria which select k to minimise

$$MIC(k) := \ln(\hat{\sigma}_k^2) + \frac{C_T(\tau(k) + k)}{T - k_{\max}} \tag{4.5}$$

where, using the seasonal notation, $\hat{\sigma}_k^2 := (T - k_{\max})^{-1} \sum_t \sum_s \widehat{e}_{4t+s}^k$, with the residuals $\widehat{e}_{4t+s}^k$ obtained from the unit root test regression augmented with k lags of the dependent variable, and the summation $\sum_t \sum_s$ extends from $4t + s = k_{\max} + 1$ to $4t + s = 4N$, while C_T is defined by the specific criterion ($C_T := 2$ for $MAIC$ and $C_T := \ln(T - k_{\max})$ for $MBIC$). Applying the approach of Ng and Perron (2001, pp.1528-1529) to the seasonal unit root null hypothesis of (2.2), and noting the asymptotic orthogonality between the integrated and stationary regressors in (4.4), as well as the mutual (asymptotic) orthogonality of the regressors $x_{j,4t+s}$, $j = 0, 1, 2$, and $x_{1,4t+s}^*$, leads to

$$\tau_T(k) := \left(\hat{\sigma}_k^2\right)^{-1} \left(\sum_{j=0}^{2} \left[\hat{\pi}_j^2 \sum_t \sum_s (x_{j,4t+s})^2 \right] + \left[(\hat{\pi}_1^*)^2 \sum_t \sum_s \left(x_{1,4t+s}^*\right)^2 \right] \right) \tag{4.6}$$

as the penalty term for the quarterly augmented HEGY regression. Therefore, our $MAIC$ and $MBIC$ results employ (4.5) in conjunction with (4.6).

In the light of the finding of Hall (1994) that using an information criterion to select the maximum lag k over $k = 0, 1, ..., k_{\max}$ may lead to size distortions in cases (such as the seasonal context) where there are "gaps" in the d_j coefficients, we propose an alternative sequential method (labelled in the tables as $SAIC$, $SBIC$, $SMAIC$ and $SMBIC$). This starts by computing the criterion for $k = k_{\max}$, with the value then computed with each individual lag $1, ..., k_{\max}$ deleted, one by one. If the criterion is improved by dropping any lag, the single lag that has the least effect on the criterion is removed from (4.4), and the procedure is repeated from this new specification. This procedure stops when no improvement in the criterion results from deleting any additional individual lag.

In addition to information criteria procedures, sequential methods based on the significance of individual lag coefficients are also examined, using 5% and 10% critical values from the standard normal distribution. Following Ng and Perron (1995) and Hall (1994), one procedure "tests down" from $k_{\max}$ to determine the maximum lag k to be employed (with no gaps); these methods are denoted as t-sq(5%) and t-sq(10%), respectively[3]. Results are also reported for the approach proposed by

[3]For this procedure, and also those suggested by Beaulieu and Miron (1993) and Rodrigues and Taylor (2004a), results

Beaulieu and Miron (1993), where (4.4) is estimated for given maximum lag order $k_{\max}$ and all lagged values with coefficients individually insignificant at the 5% or 10% level are deleted in a single step; we denote these as t-$bm(5\%)$ and t-$bm(10\%)$. Finally, the sequential method used in Rodrigues and Taylor (2004a) is employed, where at each stage the least significant of any lagged dependent variable coefficient is deleted, until all remaining coefficients are significant [t-$rt(5\%)$ and t-$rt(10\%)$].

In all cases the maximum initial lag is $k_{\max} := \lfloor \ell\,(4N/100)^{1/4} \rfloor$ with $\ell = 4$ and $\ell = 12$, as employed by Schwert (1989) and others. In practice, our focus is the latter value, which is also used by Ng and Perron (2001) for the modified information criteria. For the realistic case in applied work of $N = 50$ (100) years of data, $\ell = 12$ implies the use of a maximum augmentation lag of 14 (16) quarters, whereas $\ell = 4$ leads to 4 (5) lags being considered. It is also important to note that our choice of $k_{\max}$ ensures that k in (4.4) satisfies Assumption A.3" and hence, as discussed in Remark 9, the OLS estimator of $\Phi := [d_1, \ldots, d_k]'$ is consistent, as required for the use of data-based lag selection when the innovation process ε_{4t+s} satisfies the general Assumption A.1. In practice, however, results are shown for the case where ε_{4t+s} is an IID $N(0,1)$ sequence. Although we experimented with a variety of non-IID martingale difference specifications for ε_{4t+s}, the results are almost identical to those reported and, hence, are omitted.

Having determined the augmentation, all HEGY unit root tests employ a nominal 5% significance level, using asymptotic critical values. Results for empirical size ($c = 0$) are reported in Tables 1 to 3, with corresponding results for size-adjusted power ($c = 10$) in Tables 4 to 6. These tables relate to a selection of the cases for (4.2) and (4.3) detailed above, with sufficient results shown to indicate the general patterns.

4.2 Size Properties

Although the DGP employed for Table 1 is a seasonal random walk with IID innovations, and hence no lag augmentation is required in (4.4), the results show a number of interesting characteristics. Firstly, the parameterisation resulting from the use of modified information criteria (that is, $MAIC$, $MBIC$, $SMAIC$ and $SMBIC$) leads to under-sized tests for this case, with the sequential versions being particularly poorly sized with $\ell = 12$. On the other hand, the tests based on the significance of individual t-ratio statistics in (4.4) tend to be over-sized, with this particularly notable for t-$bm(10\%)$ and t-$rt(10\%)$ when $N = 50$ and tests are applied to all seasonal unit roots (F_{12} and F_{012}). Across all methods in Table 1, the use of more highly parameterised models, resulting from $k_{\max}$ specified using $\ell = 12$, almost always results in poorer size than $\ell = 4$. As anticipated, and although $k_{\max}$ increases with N, empirical size typically improves when the larger sample is employed.

Subsequent size results are shown only for $\ell = 12$. Not surprisingly, $\ell = 4$ results in better size than those shown when the true process is an autoregression of order less than $k_{\max}$, but it can perform very poorly when the DGP has an MA form (even when $N = 100$) or when $k_{\max}$ under-specifies the

were also obtained for a significance level of 15%. These are excluded to conserve space, but exhibit qualitatively similar patterns to the corresponding 10% ones.

true AR order.

Table 2 examines four moving average processes. An MA(1) coefficient $\theta_1 = -0.8$ (left-hand of Panel A) is fairly close to cancellation with the AR unit root -1, hence distorting inference at the Nyquist frequency (t_2, together with F_{12} and F_{012}). Indeed, with $N = 50$, BIC leads to a rejection probability of more than a half at this frequency. Since this is the situation for which modified criteria are designed and higher augmentation improves the approximation to this process, $MAIC$ performs relatively well at the Nyquist frequency across both sample sizes. Sequential lag selection ($SMAIC$) further improves this performance, with empirical sizes of 0.056 and 0.044 for $N = 50$ and 100 respectively, with $SMBIC$ having very similar size at this frequency. At other frequencies, however, tests based on (4.5) remain under-sized, as in Table 1, with the good performance of $MAIC$ and $MBIC$ for the joint tests being largely fortuitous and results from under- and over-sizing for the tests at different frequencies. Among the hypothesis testing approaches to lag selection, *t-sq*(10%) delivers the best overall size performance, despite being substantially over-sized at the Nyquist frequency, which also affects F_{12} and F_{012}. Although not shown, the mirror image case of $\theta_1 = 0.8$ interchanges the roles of the zero and Nyquist frequencies, and leads to over-sizing for t_0 as in Hall (1994) and Ng and Perron (2001) for the Dickey-Fuller test. On the other hand, the MA(2) with $\theta_1 = 0, \theta_2 = -0.64$ relates to the harmonic seasonal frequency, resulting in substantial over-sizing of F_1, together with F_{12} and F_{012}, except when lags are selected using the modified criteria (right side of Panel A, Table 1). The negatively autocorrelated MA(2) with $(1-0.64L^2) = (1+0.8L)(1-0.8L)$ (results not shown) leads to similar empirical sizes for t_2 and t_0 as those discussed for corresponding MA(1) processes with $\theta_1 = \mp 0.8$.

Recognizing that $(1-0.5L^4) = (1-0.84L)(1+0.84L)(1+0.71L^2)$, these patterns from the simple MA(1) and MA(2) processes carry over to the seasonal MA with $\Theta = 0.5$ in Panel B of Table 2, where tests at all frequencies are subject to over-sizing and the greatest distortions apply when BIC is used, while $MAIC$ has most reliable size overall. On the other hand, the seasonal MA with $\Theta = -0.5$ does not approximate any AR unit root in the DGP, and hence the patterns of Table 1 largely apply in this case, albeit with a little more distortion. Surprisingly, across Panel B of Table 2, for both $N = 50$, 100, methods that allow elimination of intermediate lags in the augmentation polynomial through hypothesis tests (*t-bm* and *t-rt*) almost always lead to poorer size performance that the corresponding *t-sq* procedure that has no gaps, despite the implied AR approximation having a seasonal form. Although $SAIC$ and $SBIC$ improve on AIC and BIC, respectively when $\Theta = 0.5$, their advantages are less evident when $\Theta = -0.5$, while there is also little indication in Panel B that sequential lag elimination aids the modified information criteria. However, although not shown, it may be noted that (analogously to the Nyquist frequency results in Panel A) $SMAIC$ and $SMBIC$ have better size performance than other methods across all frequencies for a seasonal MA process with $\Theta = 0.7$, where the near-cancellation of the (zero and seasonal) AR unit roots with the corresponding MA components applies even more strongly than when $\Theta = 0.5$.

The AR processes of Table 3 provide further evidence that intermediate lag elimination can increase size distortions, even when the true AR polynomials have some intermediate zero coefficients; this now applies across $SAIC$, $SBIC$, $SMAIC$, $SMBIC$, *t-bm* and *t-rt* in both Panels A and B, in comparison with the corresponding procedures with no such intermediate lag elimination. An interesting feature of Table 3 is that while size usually improves with N (as anticipated), this does not apply here for the modified information criteria ($MAIC$ and $MBIC$) and their sequential variants, which are always under-sized.

Overall, the results across Tables 1-3 indicate that (with quarterly data), no single procedure leads to reliable size performance across all cases when autocorrelation may be present. In line with the arguments of Ng and Perron (2001) in the context of a conventional unit root, we find that the seasonal unit root versions of $MAIC$ and (to a slightly lesser extent) $MBIC$ deliver relatively good size for tests at frequencies when an MA disturbance process has root(s) that are close to cancelling with the corresponding AR unit root(s). Indeed, the sequential versions, $SMAIC$ and $SMBIC$, perform very well in this respect. In other cases, however, they (and especially the sequential versions) are under-sized. Although it has good size for AR processes (providing $k_{\max}$ contains sufficient lags), conventional AIC can be very poorly sized for MA processes, as shown by Ng and Perron (2001) for Dickey-Fuller tests. In general, AIC dominates SIC. The "testing down" procedure *t-sq* delivers quite good size for AR processes, and is better than AIC for moving averages. While sequential intermediate lag elimination can sometimes aid information criteria approaches, such elimination based on hypothesis tests typically delivers less reliable size that *t-sq* without gaps, even when the DGP implies some zero coefficients in the augmentation lag polynomial. However, it needs to be emphasized that all procedures examined in detail here allow the same maximum lag of $\lfloor 12\,(4N/100)^{1/4} \rfloor$, equating to 14 and 16 quarters for $N = 50$ or 100 years of data. Any procedure that starts from a low maximum lag of (say) 4 or 5 quarters can have poor size in the presence of MA disturbances (even of low order) or an AR process of order higher than $k_{\max}$.

4.3 Size-Adjusted Power Properties

Size-adjusted (local) power results are shown in Tables 4 to 6. These analyze cases corresponding to Tables 1 to 3, except that now $c = 10$ in the DGP of (4.1). The IID disturbance case of Table 4 shows, as anticipated, that more parsimonious models yield higher power. In particular, across all statistics, specifications using $\ell = 4$ almost always have power at least as great as the corresponding test based on $\ell = 12$, with the difference sometimes being fairly substantial (such as the joint F_{12} and F_{012} tests for $SMAIC$ and $SMBIC$). BIC-based lag selection also generally leads to tests with higher power that the corresponding AIC procedure, although the differences are generally modest. Similarly, the t-test lag selection procedures yield marginally higher power when a tighter significance level is used. Note, however, that power is largely constant for $N = 50$, 100, because a local-to-unity DGP is employed in both cases.

Power is typically fairly modest in Table 4 for the tests t_0 and t_2, which is explained by the AR polynomial being $(1 - 0.8L^4) = (1 - 0.95L)(1 + 0.95L)(1 + 0.89L^2)$ for $c = 10$ and $N = 50$, indicating the proximity to unit roots, especially at the zero and Nyquist frequencies. This continues to apply for the processes in Table 5, Panel A. When near-cancellation occurs across the MA and AR components (namely the t_2 statistic when $\theta_1 = -0.8$ and F_1 with $\theta_2 = -0.64$), the conventional AIC criterion delivers relatively good power, with $SBIC$ being comparable for the MA(2) case. However, the Beaulieu and Miron (1993) approach not only yields similar size-adjusted power to AIC at these frequencies, but it also performs well relative to other procedures across all frequencies, with t-rt typically having power intermediate between t-sq and t-bm. Although BIC often has higher power than t-bm for the annual processes of Panel B, nevertheless the latter continues to show the highest power overall among the hypothesis testing approaches. It is also noteworthy that while $MAIC$ and $MBIC$, together with their sequential versions, do not generally suffer severe power losses compared to AIC and BIC for the processes of Panel A, they can be very poor for the processes in Panel B when $\Theta = -0.5$.

For the AR processes of Table 6, all of which have a seasonal form, intermediate lag deletion, whether based on information criteria or hypothesis testing, can be beneficial for the power of tests at individual zero or seasonal frequencies in the presence of positive autocorrelation (left-hand of Panels A and B), although this is less evident for the overall F_{012} test. However, all procedures have lower power when $\Phi = -0.5$ compared with $\Phi = 0.5$. Analogously to results of Ng and Perron (1995) for the ADF test, the modified information criteria methods ($MAIC$, $MBIC$, $SMAIC$, $SBIC$) have particularly poor power for $\Phi = -0.5$ when $N = 50$; this also applies at $N = 100$ for $MBIC$ and $SBIC$. Comparing power across Panels A and B, it is further evident that a negative AR(1) coefficient, in addition to the seasonal coefficient $\Phi = -0.5$, further reduces power across all methods, with the partial exception of the MIC procedures with low power in Panel A.

From the perspective of size-adjusted power, therefore, the Beaulieu and Miron (1993) approach to lag selection performs relatively well across both the MA and AR error processes considered, with this also applying to AIC. However, these procedures can have relatively poor size for MA processes (Table 2). The procedure least prone to size distortions overall in Tables 1-3 is $MAIC$, followed by t-sq(10%). Although the latter typically has power only modestly less than that of t-bm, $MAIC$ can have relatively poor power for error processes that have a positively autocorrelated seasonal MA form or a seasonal AR form with negative coefficients.

5 Conclusions

This paper has made two key contributions to the literature on regression-based seasonal unit root testing. First, we have extended the results relating to the asymptotic null distribution for the ADF unit root tests given in Chang and Park (2002) to the case of augmented HEGY seasonal unit root tests. Specifically, we have shown that regression t-statistics for unit roots at the zero and Nyquist frequencies and all F-type statistics have pivotal limiting null distributions in the case where the shocks follow a general linear process driven by martingale difference innovations, but that this is not the case for the t-statistics at the harmonic seasonal frequencies whose asymptotic null distributions depend on serial correlation nuisance parameters. The rate at which the length of the lag augmentation polynomial used in the test regression is required to increase for these results to hold was also explored and shown to coincide with the rate derived for the non-seasonal ADF statistic by Chang and Park (2002).

Second, through Monte Carlo simulation experiments, we have explored the performance of a variety of popularly employed data-based methods for determining the lag augmentation polynomial in the HEGY test regression. Further, we have extended the modified information criteria approach of Ng and Perron (2001) to the seasonal unit root testing context, which is implemented as $MAIC$ and $MBIC$. In general, the procedure of Beaulieu and Miron (1993) applied using a 10% significance level, performs quite well, typically not exhibiting severe size distortions and being competitive on power. Deleting intermediate lags can be advantageous in terms of power, but such deletion may also increase size distortions, especially in the presence of MA disturbances. Information criteria approaches can have poor size (for example, AIC/BIC with MAs) or poor power ($MAIC/MBIC$ for disturbance processes with negative seasonal AR coefficients). If size is the most important consideration, the seasonal $MAIC$ may be recommended, with our sequential lag elimination version $SMAIC$ delivering good size when near-cancellation occurs across AR unit roots with corresponding roots in the MA disturbance process.

References

Beaulieu, J.J. and Miron, J.A., (1993) Seasonal unit roots in aggregate U.S. data. Journal of Econometrics 55, 305-328.

Berk, K.N., (1974) Consistent autoregressive spectral estimates. The Annals of Statistics 2, 389-502.

Boswijk, H.P., and Franses, P.H., (1996). Unit roots in periodic autoregressions. Journal of Time Series Analysis 17, 221-245.

Box, G.E.P., and Jenkins, G.M., (1976) Time Series Analysis: Forecasting and Control (revised edition). San Francisco: Holden-Day.

Burridge, P. and Taylor, A. M. R., (2001). On the properties of regression-based tests for seasonal unit roots in the presence of higher-order serial correlation. Journal of Business and Economic Statistics 3, 374-379.

Chang, Y. and Park, J.Y., (2002). On the asymptotics of ADF tests for unit roots. Econometric Reviews 21, 431-447.

Davis, P.J., (1979). Circulant Matrices. Wiley-Interscience: New York.

Davidson, J., (1994). Stochastic Limit Theory. Oxford University Press: Oxford.

del Barrio Castro, T., (2007). Using the HEGY procedure when not all the roots are present. Journal of Time Series Analysis 29, 910-922.

del Barrio Castro T. and Osborn D.R., (2011). HEGY tests in the presence of moving averages. Oxford Bulletin of Economic and Statistics, forthcoming.

Fuller, W.A., (1996). Introduction to Statistical Time Series, Second Edition. Wiley: New York.

Ghysels, E., Lee, H. S., and Noh, J., (1994). Testing for unit roots in seasonal time series: some theoretical extensions and a Monte Carlo investigation. Journal of Econometrics 62, 415- 442,

Ghysels, E., Osborn, D.R., (2001). The Econometric Analysis of Seasonal Time Series. Cambridge University Press: Cambridge.

Gray, R. M., (2006). Toeplitz and Circulant Matrices, A Review. Foundation and Trends(R) in Communications and Information Theory: Now Publishers Inc.

Hall, A.R., (1994). Testing for a unit root in time series with pretest data-based model selection. Journal of Business and Economic Statistics 12, 461-70.

Hamilton J., (1994). Time Series Analysis. Princeton University Press: Princeton.

Hylleberg, S., Engle, R. F., Granger, C. W. J., and Yoo, B. S., (1990). Seasonal integration and cointegration. Journal of Econometrics 44, 215-238.

Ng, S., and Perron, P., (1995). Unit root tests in ARMA models with data dependent methods for selection of the truncation lag. Journal of the American Statistical Association 90, 268-281.

Ng, S., and Perron, P., (2001). Lag length selection and the construction of unit root tests with good size and power. Econometrica 69, 1519-1554.

Osborn, D. R., and Rodrigues, P. M. M., (2002). Asymptotic distributions of seasonal unit root tests: a unifying approach. Econometric Reviews 21, 221-241.

Phillips, P.C.B., and Durlauf, S.N., (1986). Multiple time series regression with integrated processes. Review of Economic Studies 53, 473-95.

Rodrigues, P. M. M., and Taylor A. M. R., (2004a). Alternative estimators and unit root tests for seasonal autoregressive processes. Journal of Econometrics 120, 35-73.

Rodrigues, P. M. M., and Taylor A. M. R., (2004b). Asymptotic distributions for regression-based seasonal unit root test statistics in a near-integrated model. Econometric Theory 20, 645-670.

Said, S.E., and Dickey, D. A., (1984). Testing for unit roots in autoregressive-moving average models of unknown order. Biometrika 71, 599-607.

Schwert, G., (1989). Testing for unit roots: a Monte Carlo investigation. Journal of Business and Economic Statistics 7, 147-158.

Smith R.J., and Taylor, A.M.R., (1998). Additional critical values and asymptotic representations for seasonal unit root tests. Journal of Econometrics 85, 269-288.

Smith, R.J., and Taylor, A.M.R., (1999). Likelihood ratio tests for seasonal unit roots. Journal of Time Series Analysis 20, 453-476.

Smith R.J., Taylor, A.M.R., and del Barrio Castro, T., (2009). Regression-based seasonal unit root tests. Econometric Theory 25, 527-560.

Taylor, A. M. R., (1997). On the practical problems of computing seasonal unit root tests. International Journal of Forecasting 13, 307-318.

Taylor, A. M. R., (2005). Variance ratio tests of the seasonal unit root hypothesis. Journal of Econometrics 124, 33-54.

A Appendix

For the purposes of the Appendix, and without loss of generality, we simplify our exposition by setting the initial vector $X_0 := (x_{1-S}, ..., x_0)' = \mathbf{0}$ throughout.

Proof of Lemma 1: The conditions placed on $\{E_t\}$ through Assumption A.1 imply that it satisfies a multivariate invariance principle, see for example Phillips and Durlauf (1996) or Davidson (1994, pp.454-455), such that

$$N^{-1/2} \sum_{j=1}^{\lfloor rN \rfloor} E_j \xrightarrow{d} \sigma \mathbf{W}(r) \tag{A.1}$$

where $\mathbf{W}(r)$ is a $S \times 1$ vector standard Brownian motion. Next observe from (3.1) and (3.2), that

$$\begin{aligned} N^{-1/2} X_{\lfloor rN \rfloor} &= N^{-1/2} \sum_{j=1}^{\lfloor rN \rfloor} U_j \\ &= \mathbf{\Psi}(1) N^{-1/2} \sum_{j=1}^{\lfloor rN \rfloor} E_j + o_p(1) \end{aligned}$$

where the approximation in the second line follows from the same argument as in Boswijk and Franses (1996, p.238). Using (A.1) and the continuous mapping theorem [CMT] the result in (3.3) then follows immediately.

To establish the second part of the lemma, observe first that $I_S = \frac{1}{S} C_0 + \frac{1}{S} C_{S/2} + \frac{2}{S} \sum_{j=1}^{S^*} C_j$. Moreover, noting that $\mathbf{\Psi}(1)$ is also a circulant matrix, then by the properties of products of circulant matrices it can be shown that $C_0 \mathbf{\Psi}(1) = \psi(1) C_0$, $C_{S/2} \mathbf{\Psi}(1) = \psi(-1) C_{S/2}$, $C_j \mathbf{\Psi}(1) = b_j C_j + a_j C_j^*$ and $C_j^* \mathbf{\Psi}(1) = -a_j C_j + b_j C_j^*$ for $j = 1, \ldots S^*$; see, *inter alia*, Theorem 3.2.4 of Davis (1979), Theorem 3.1 of Gray (2006) and Smith *et al.* (2009) for further details. The stated result then follows immediately.

Proof of Lemma 2: Noting that $X_{j,t} = C_j X_t$, $j = 0, \ldots, \lfloor S/2 \rfloor$, and that $X_{i,t}^* = C_i^* X_t$, $i = 1, \ldots, S^*$, the stated results follow immediately from Lemma 1, using the following identities: $C_0 C_0 = S C_0$, $C_{S/2} C_{S/2} = S C_{S/2}$, $C_j C_j = \frac{S}{2} C_j$, $C_j C_j^* = \frac{S}{2} C_j^*$ and $C_j^* C_j^* = \frac{S}{2} C_j$, $j = 1, ..., S^*$, also recalling from Remark 4 that the remaining matrix products between C_0, $C_{S/2}$, C_j and C_j^*, $j = 1, \ldots, S^*$ are all zero matrices, and noting that multiplication between circulant matrices is commutative.

For later reference, noting that $C_0 = \mathbf{v}_0 \mathbf{v}_0{}'$, where $\mathbf{v}_0{}' = [1, 1, 1, ..., 1]$, $C_{S/2} = \mathbf{v}_{S/2} \mathbf{v}_{S/2}{}'$, where $\mathbf{v}_{S/2}{}' = [-1, 1, -1, ..., 1]$, and that $C_j = \mathbf{v}_j \mathbf{v}_j'$ and $C_j^* = \mathbf{v}_j \mathbf{v}_j^{*\prime}$, where

$$\mathbf{v}_j' = \begin{bmatrix} \cos(\omega_j [1-S]) & \cos(\omega_j [2-S]) & \cdots & \cos(0) \\ \sin(\omega_j [1-S]) & \sin(\omega_j [2-S]) & \cdots & \sin(0) \end{bmatrix}$$

and

$$\mathbf{v}_j^{*\prime} = \begin{bmatrix} -\sin(\omega_j [1-S]) & -\sin(\omega_j [2-S]) & \cdots & -\sin(0) \\ \cos(\omega_j [1-S]) & \cos(\omega_j [2-S]) & \cdots & \cos(0) \end{bmatrix},$$

it is straightforwardly seen that

$$N^{-1/2}x_{0,S\lfloor rN\rfloor+s} = \frac{\psi(1)}{\sqrt{N}}\left(\sum_{h=1}^{S}\left[\sum_{j=1}^{\lfloor rN\rfloor}\varepsilon_{Sj-(S-h)}\right]\right)+o_p(1) \tag{A.2}$$

$$N^{-1/2}x_{S/2,S\lfloor rN\rfloor+s} = \frac{\psi(-1)(-1)^s}{\sqrt{N}}\left(\sum_{h=1}^{S}\left[(-1)^h\sum_{j=1}^{\lfloor rN\rfloor}\varepsilon_{Sj-(S-h)}\right]\right)+o_p(1) \tag{A.3}$$

and that for $i=1,...,S^*$,

$$\begin{aligned} N^{-1/2}x_{i,S\lfloor rN\rfloor+s} &= \frac{b_i}{\sqrt{N}}\left[\cos(\omega_i[s])\left(\sum_{h=1}^{S}\cos(\omega_i[h-S])\left[\sum_{j=1}^{\lfloor rN\rfloor}\varepsilon_{Sj-(s-h)}\right]\right)\right. \\ &\left.+\sin(\omega_i[s])\left(\sum_{h=1}^{S}\sin(\omega_i[h-S])\left[\sum_{j=1}^{\lfloor rN\rfloor}\varepsilon_{Sj-(s-h)}\right]\right)\right] \\ &+\frac{a_i}{\sqrt{N}}\left[\sin(\omega_i[s])\left(\sum_{h=1}^{S}\cos(\omega_i[h-S])\left[\sum_{j=1}^{\lfloor rN\rfloor}\varepsilon_{Sj-(s-h)}\right]\right)\right. \\ &\left.-\cos(\omega_i[s])\left(\sum_{h=1}^{S}\sin(\omega_i[h-S])\left[\sum_{j=1}^{\lfloor rN\rfloor}\varepsilon_{Sj-(s-h)}\right]\right)\right]+o_p(1) \end{aligned} \tag{A.4}$$

and

$$\begin{aligned} N^{-1/2}x^*_{i,S\lfloor rN\rfloor+s} &= \frac{b_i}{\sqrt{N}}\left[\sin(\omega_i[s])\left(\sum_{h=1}^{S}\cos(\omega_i[h-S])\left[\sum_{j=1}^{\lfloor rN\rfloor}\varepsilon_{Sj-(s-h)}\right]\right)\right. \\ &\left.-\cos(\omega_i[s])\left(\sum_{h=1}^{S}\sin(\omega_i[h-S])\left[\sum_{j=1}^{\lfloor rN\rfloor}\varepsilon_{Sj-(s-h)}\right]\right)\right] \\ &-\frac{a_i}{\sqrt{N}}\left[\cos(\omega_i[s])\left(\sum_{h=1}^{S}\cos(\omega_i[h-S])\left[\sum_{j=1}^{\lfloor rN\rfloor}\varepsilon_{Sj-(s-h)}\right]\right)\right. \\ &\left.+\sin(\omega_i[s])\left(\sum_{h=1}^{S}\sin(\omega_i[h-S])\left[\sum_{j=1}^{\lfloor rN\rfloor}\varepsilon_{Sj-(s-h)}\right]\right)\right]+o_p(1). \end{aligned} \tag{A.5}$$

Proof of Proposition 1: First re-write (2.9) in vector form, viz,

$$\mathbf{y}=(\mathbf{Y},\mathbf{Z}_k)\boldsymbol{\beta}_0+\mathbf{u}$$

where $\mathbf{y}$ is a $T\times 1$ vector with generic element $\Delta_S x_{St+s}$; $\mathbf{Y} := [\mathbf{y}_0|\mathbf{y}_1|\mathbf{y}_1^*|\mathbf{y}_2|\mathbf{y}_2^*|\ldots\mathbf{y}_{S^*}|\mathbf{y}_{S^*}^*|\mathbf{y}_{S/2}]$ is a $T\times S$ matrix where $\mathbf{y}_i$, $i=0,...,\lfloor S/2\rfloor$ are $T\times 1$ vectors with generic element $x_{i,St+s}$, and $\mathbf{y}_i^*$, $i=1,...,S^*$ are $T\times 1$ vectors with generic element $x^*_{i,St+s}$; $\mathbf{Z}_k := [\mathbf{z}_1|\mathbf{z}_2|\cdots|\mathbf{z}_k]$ is a $T\times k$ matrix with $\mathbf{z}_j$ being $T\times 1$ vectors with generic element $\Delta_S x_{St+s-j}$ for $j=1,\ldots,k$; $\boldsymbol{\beta}_0=[\Pi'|\Phi']'$ $:= [\pi_0,\pi_1,\pi_1^*,\pi_2,\pi_2^*,\ldots,\pi_{S^*},\pi_{S^*}^*,\pi_{S/2},,d_1,\ldots,d_k]'$; finally $\mathbf{u}$ is a $T\times 1$ vector with generic element e^k_{St+s}. Commensurate with the partitions of $\boldsymbol{\beta}_0$, define the $(S+k)\times(S+k)$ scaling matrix, $M :=$

diag $[T, \ldots, T, T^{1/2}, \ldots, T^{1/2}]$. It is then straightforwardly seen that the OLS estimator, $\hat{\beta}_0$ say, from (2.9) is such that

$$M\left(\hat{\beta}_0 - \beta_0\right) = \begin{bmatrix} T^{-2}\mathbf{Y}'\mathbf{Y} & T^{-3/2}\mathbf{Y}'\mathbf{Z}_k \\ T^{-3/2}\mathbf{Z}_k'\mathbf{Y} & T^{-1}\mathbf{Z}_k'\mathbf{Z}_k \end{bmatrix}^{-1} \times \begin{bmatrix} T^{-1}\mathbf{Y}'\mathbf{u} \\ T^{-1/2}\mathbf{Z}_k'\mathbf{u} \end{bmatrix}. \tag{A.6}$$

Due to the unit root non-stationary of the elements of $\mathbf{Y}$ and the stationarity of the elements of $\mathbf{Z}_k$, it follows immediately that $T^{-3/2}\mathbf{Y}'\mathbf{Z}_k \xrightarrow{p} \mathbf{0}$, so that the inverse matrix in (A.6) is asymptotically block diagonal. As a consequence, the scaled estimators $T\left(\hat{\Pi} - \Pi\right)$ and $T^{1/2}\left(\hat{\Phi} - \Phi\right)$ are asymptotically orthogonal. Moreover, it is straightforward to show that $T^{-2}\mathbf{Y}'\mathbf{Y}$ weakly converges to a $S \times S$ diagonal matrix. We may therefore consider the large sample behaviour of the OLS estimators of π_j, $j = 0, ..., \lfloor S/2 \rfloor$, and π_i^*, $i = 1, ..., S^*$, separately.

Consequently, defining $\mathbf{Q} := I_T - \mathbf{Z}_k\left(\mathbf{Z}_k'\mathbf{Z}_k\right)^{-1}\mathbf{Z}_k'$, we can write the so-called *normalized bias* statistics as follows:

$$\begin{aligned} T\widehat{\pi}_j &= \frac{T^{-1}\mathbf{y}_j'\mathbf{Q}\mathbf{u}}{T^{-2}\mathbf{y}_j'\mathbf{Q}\mathbf{y}_j} + o_p(1) \\ &= \frac{T^{-1}\left(\mathbf{y}_j'\mathbf{u} - \mathbf{y}_j'\mathbf{Z}_k\left(\mathbf{Z}_k'\mathbf{Z}_k\right)^{-1}\mathbf{Z}_k'\mathbf{u}\right)}{T^{-2}\left(\mathbf{y}_j'\mathbf{y}_j - \mathbf{y}_j'\mathbf{Z}_k\left(\mathbf{Z}_k'\mathbf{Z}_k\right)^{-1}\mathbf{Z}_k'\mathbf{y}_j\right)} + o_p(1), \quad j = 0, ..., \lfloor S/2 \rfloor \end{aligned} \tag{A.7}$$

and

$$\begin{aligned} T\widehat{\pi}_i^* &= \frac{T^{-1}\mathbf{y}_i^{*\prime}\mathbf{Q}\mathbf{u}}{T^{-2}\mathbf{y}_i^{*\prime}\mathbf{Q}\mathbf{y}_i^*} + o_p(1) \\ &= \frac{T^{-1}\left(\mathbf{y}_i^{*\prime}\mathbf{u} - \mathbf{y}_i^{*\prime}\mathbf{Z}_k\left(\mathbf{Z}_k'\mathbf{Z}_k\right)^{-1}\mathbf{Z}_k'\mathbf{u}\right)}{T^{-2}\left(\mathbf{y}_i^{*\prime}\mathbf{y}_i^* - \mathbf{y}_i^{*\prime}\mathbf{Z}_k\left(\mathbf{Z}_k'\mathbf{Z}_k\right)^{-1}\mathbf{Z}_k'\mathbf{y}_j^*\right)} + o_p(1), \quad j = 1, ..., S^*. \end{aligned} \tag{A.8}$$

The following lemma will allow us to simplify the expressions given above.

Lemma A.1 *Under the conditions of Proposition 1 the following results hold: (i)* $\|\mathbf{Z}_k'\mathbf{Z}_k^{-1}\| = O_p\left(T^{-1}\right)$*; (ii)* $\|\mathbf{Z}_k'\mathbf{u}\| = o_p\left(Tk^{-1/2}\right)$*, and (iii)* $\|\mathbf{Z}_k'\mathbf{y}_j\| = O_p\left(Tk^{1/2}\right)$*,* $j = 0, ..., \lfloor S/2 \rfloor$ *and* $\|\mathbf{Z}_k'\mathbf{y}_h^*\| = O_p\left(Tk^{1/2}\right)$*,* $h = 1, ..., S^*$*.*

Proof: Parts (i) and (ii) follow immediately from parts (a) and (c) respectively of Lemma 3.2 of Chang and Park (2002) simply by replacing the standard first differences which appear there by the seasonal differences, $\Delta_S x_{St+s}$. In order to establish the validity of part (iii), first use the fact that for $j = 0, ..., \lfloor S/2 \rfloor$, $\|\mathbf{Z}_k'\mathbf{y}_j\|^2 = \sum_{i=1}^k \left(\mathbf{z}_i'\mathbf{y}_j\right)^2$, where $\mathbf{z}_i'\mathbf{y}_j = \sum_{t=1}^N \sum_{s=1-S}^0 \Delta_S x_{St+s-i} x_{j,St+s}$, $i = 1, ..., k$, and, for $h = 1, ..., S^*$, $\|\mathbf{Z}_k'\mathbf{y}_h^*\|^2 = \sum_{i=1}^k \left(\mathbf{z}_i'\mathbf{y}_h^*\right)^2$, where $\mathbf{z}_i'\mathbf{y}_h^* = \sum_{t=1}^N \sum_{s=1-S}^0 \Delta_S x_{St+s-i} x_{h,St+s}^*$ $i = 1, \ldots, k$. Next observe that the following equalities hold for each $i = 1, ..., k$:

$$\sum_{t=1}^N \sum_{s=1-S}^0 x_{0,St+s}\Delta_S x_{St+s-i} = \sum_{t=1}^N \sum_{s=1-S}^0 \left(\sum_{j=1}^{(St+s)-1} \Delta_S x_{St+s-j}\right)\Delta_S x_{St+s-i} \tag{A.9}$$

$$\sum_{t=1}^N \sum_{s=1-S}^0 x_{S/2,St+s}\Delta_S x_{St+s-i} = \sum_{t=1}^N \sum_{s=1-S}^0 \left(\sum_{j=1}^{(St+s)-1} (-1)^j \Delta_S x_{St+s-j}\right)\Delta_S x_{St+s-i} \tag{A.10}$$

and, moreover, that for each $h = 1, ..., S^*$, the following equalities also hold for each $i = 1, ..., k$:

$$\sum_{t=1}^{N}\sum_{s=1-S}^{0} x_{h,St+s}\Delta_S x_{St+s-i} = \sum_{t=1}^{N}\sum_{s=1-S}^{0}\left(\sum_{j=1}^{(St+s)-1}\cos(j\omega_h)\Delta_S x_{St+s-j}\right)\Delta_S x_{St+s-i} \quad \text{(A.11)}$$

$$\sum_{t=1}^{N}\sum_{s=1-S}^{0} x^*_{h,St+s}\Delta_S x_{St+s-i} = \sum_{t=1}^{N}\sum_{s=1-S}^{0}\left(\sum_{j=1}^{(St+s)-1}-\sin(j\omega_h)\Delta_S x_{St+s-j}\right)\Delta_S x_{St+s-i}. \quad \text{(A.12)}$$

The stated result is then established using (A.9)-(A.12) and following along the lines of the proof of part (b) of Lemma 3.2 in Chang and Park (2002, pp.443-444). □

Applying the results in Lemma A.1 to (A.7) and (A.8), we then have that

$$\begin{aligned} T\widehat{\pi}_j &= \frac{T^{-1}\mathbf{y}_j'\mathbf{u}}{T^{-2}\mathbf{y}_j'\mathbf{y}_j} + o_p(1) \\ &= \frac{T^{-1}\sum_{t=1}^{N}\sum_{s=1-S}^{0} x_{j,St+s}e^k_{St+s}}{T^{-2}\sum_{t=1}^{N}\sum_{s=1-S}^{0} x^2_{j,St+s}} + o_p(1), \quad j = 0, ..., \lfloor S/2 \rfloor \end{aligned} \quad \text{(A.13)}$$

and

$$\begin{aligned} T\widehat{\pi}^*_h &= \frac{T^{-1}\mathbf{y}_h^{*\prime}\mathbf{u}}{T^{-2}\mathbf{y}_h^{*\prime}\mathbf{y}^*_h} + o_p(1) \\ &= \frac{T^{-1}\sum_{t=1}^{N}\sum_{s=1-S}^{0} x^*_{h,St+s}e^k_{St+s}}{T^{-2}\sum_{t=1}^{N}\sum_{s=1-S}^{0} x^{*2}_{h,St+s}} + o_p(1), \quad h = 1, ..., S^*. \end{aligned} \quad \text{(A.14)}$$

For each of $j = 0, ..., \lfloor S/2 \rfloor$ in (A.13) and $h = 1, ..., S^*$ in (A.14), re-write the numerators of (A.13) and (A.14), respectively, as

$$\begin{aligned} T^{-1}\sum_{t=1}^{N}\sum_{s=1-S}^{0} x_{j,St+s}e^k_{St+s} &= T^{-1}\sum_{t=1}^{N}\sum_{s=1-S}^{0} x_{j,St+s}\varepsilon_{St+s} + T^{-1}\sum_{t=1}^{N}\sum_{s=1-S}^{0} x_{j,St+s}\left(e^k_{St+s} - \varepsilon_{St+s}\right) \\ T^{-1}\sum_{t=1}^{N}\sum_{s=1-S}^{0} x^*_{h,St+s}e^k_{St+s} &= T^{-1}\sum_{t=1}^{N}\sum_{s=1-S}^{0} x^*_{h,St+s}\varepsilon_{St+s} + T^{-1}\sum_{t=1}^{N}\sum_{s=1-S}^{0} x^*_{h,St+s}\left(e^k_{St+s} - \varepsilon_{St+s}\right). \end{aligned}$$

Substituting into (A.13) and (A.14), respectively, we therefore obtain that for, $j = 0, ..., \lfloor S/2 \rfloor$,

$$T\widehat{\pi}_j = \frac{T^{-1}\sum_{t=1}^{N}\sum_{s=1-S}^{0} x_{j,St+s}\varepsilon_{St+s} + T^{-1}\sum_{t=1}^{N}\sum_{s=1-S}^{0} x_{j,St+s}\left(e^k_{St+s} - \varepsilon_{St+s}\right)}{T^{-2}\sum_{t=1}^{N}\sum_{s=1-S}^{0} x^2_{j,St+s}} + o_p(1) \quad \text{(A.15)}$$

and that, for $h = 1, ..., S^*$,

$$T\widehat{\pi}^*_h = \frac{T^{-1}\sum_{t=1}^{N}\sum_{s=1-S}^{0} x^*_{h,St+s}\varepsilon_{St+s} + T^{-1}\sum_{t=1}^{N}\sum_{s=1-S}^{0} x^*_{h,St+s}\left(e^k_{St+s} - \varepsilon_{St+s}\right)}{T^{-2}\sum_{t=1}^{N}\sum_{s=1-S}^{0} x^{*2}_{h,St+s}} + o_p(1). \quad \text{(A.16)}$$

Consider first the denominators of (A.15) and (A.16). Using the results $\mathbf{\Psi}(1)' C_0 \mathbf{\Psi}(1) = \psi(1)^2 C_0$, $\mathbf{\Psi}(1)' C_{S/2} \mathbf{\Psi}(1) = \psi(-1)^2 C_{S/2}$ and $\mathbf{\Psi}(1)' C_j \mathbf{\Psi}(1) = \left(a_j^2 + b_j^2\right) C_j$, for $j = 1, ..., S^*$ from the multivariate invariance principle in (3.3) and the CMT we obtain that

$$\begin{aligned} T^{-2} \sum_{t=1}^{N} \sum_{s=1-S}^{0} x_{j,St+s}^2 &= T^{-2} \sum_{t=1}^{N} S \left(X_{t-1}' C_j X_{t-1}\right) + o_p(1) \qquad j = 0, S/2 \\ &\xrightarrow{d} \frac{\sigma^2}{S} \int_0^1 \mathbf{W}(r)' \, \mathbf{\Psi}(1)' C_j \, \mathbf{\Psi}(1) \, \mathbf{W}(r) \, dr \\ &= \begin{cases} \sigma^2 \psi(1)^2 \int_0^1 \mathbf{W}^*(r)' C_0 \mathbf{W}^*(r) \, dr & j = 0 \\ \sigma^2 \psi(-1)^2 \int_0^1 \mathbf{W}^*(r)' C_{S/2} \mathbf{W}^*(r) \, dr & j = S/2 \end{cases} \end{aligned} \tag{A.17}$$

where $\mathbf{W}^*(r) := \frac{1}{\sqrt{S}} \mathbf{W}(r)$ and

$$\begin{aligned} T^{-2} \sum_{t=1}^{N} \sum_{s=1-S}^{0} x_{j,St+s}^2 &= T^{-2} \sum_{t=1}^{N} \left(\frac{S}{2}\right) \left(X_{t-1}' C_j \, X_{t-1}\right) + o_p(1) \qquad j = 1, \ldots, S^* \\ &\xrightarrow{d} \frac{\sigma^2}{S^2} \left(\frac{S}{2}\right) \int_0^1 \mathbf{W}(r)' \, \mathbf{\Psi}(1)' C_j \, \mathbf{\Psi}(1) \, \mathbf{W}(r) \, dr \\ &= \frac{\sigma^2 \left(a_j^2 + b_j^2\right)}{4} \int_0^1 \mathbf{W}^\dagger(r)' C_j \mathbf{W}^\dagger(r) \, dr \end{aligned} \tag{A.18}$$

where $\mathbf{W}^\dagger(r) := \frac{1}{\sqrt{S/2}} \mathbf{W}(r)$.

Consider next the numerators of (A.15) and (A.16). In each case, for the first term it is straightforward to show that

$$\begin{aligned} T^{-1} \sum_{t=1}^{N} \sum_{s=1-S}^{0} x_{j,St+s} \varepsilon_{St+s} &= T^{-1} \sum_{t=1}^{N} X_{t-1}' C_j E_t + o_p(1), \quad j = 0, \ldots, \lfloor S/2 \rfloor \\ T^{-1} \sum_{t=1}^{N} \sum_{s=1-S}^{0} x_{h,St+s}^* \varepsilon_{St+s} &= T^{-1} \sum_{t=1}^{N} X_{t-1}' C_h^* E_t + o_p(1) \quad h = 1, \ldots, S^*. \end{aligned}$$

Again using (3.3), applications of the CMT and the identities: $\mathbf{\Psi}(1)' C_0 = \psi(1) C_0$, $\mathbf{\Psi}(1)' C_{S/2} = \psi(-1) C_{S/2}$, and, for $j = 1, ..., S^*$, $\mathbf{\Psi}(1)' C_j = b_j C_j - a_j C_j^*$ and $\mathbf{\Psi}(1)' C_j^* = a_j C_j + b_j C_j^*$, the following results then obtain

$$\begin{aligned} T^{-1} \sum_{t=1}^{N} \sum_{s=1-S}^{0} x_{0,St+s} \varepsilon_{St+s} &\xrightarrow{d} \frac{\sigma^2}{S} \int_0^1 \mathbf{W}(r)' \, \mathbf{\Psi}(1)' C_0 d\mathbf{W}(r) \\ &= \sigma^2 \psi(1) \int_0^1 \mathbf{W}^*(r)' C_0 d\mathbf{W}^*(r) \end{aligned} \tag{A.19}$$

$$\begin{aligned} T^{-1} \sum_{t=1}^{N} \sum_{s=1-S}^{0} x_{S/2,St+s} \varepsilon_{St+s} &\xrightarrow{d} \frac{\sigma^2}{S} \int_0^1 \mathbf{W}(r)' \, \mathbf{\Psi}(1)' C_{S/2} d\mathbf{W}(r) \\ &= \sigma^2 \psi(-1) \int_0^1 \mathbf{W}^*(r)' C_{S/2} d\mathbf{W}^*(r) \end{aligned} \tag{A.20}$$

and for $j = 1, ..., S^*$,

$$T^{-1}\sum_{t=1}^{N}\sum_{s=1-S}^{0} x_{j,St+s}\varepsilon_{St+s} \xrightarrow{d} \frac{\sigma^2}{S}\int_0^1 \mathbf{W}(r)'\,\mathbf{\Psi}(1)'\,C_j d\mathbf{W}(r) \tag{A.21}$$

$$= \frac{\sigma^2 b_j}{2}\int_0^1 \mathbf{W}^{\dagger}(r)'\,C_j d\mathbf{W}^{\dagger}(r) - \frac{\sigma^2 a_j}{2}\int_0^1 \mathbf{W}^{\dagger}(r)'\,C_j^* d\mathbf{W}^{\dagger}(r)$$

$$T^{-1}\sum_{t=1}^{N}\sum_{s=1-S}^{0} x^*_{j,St+s}\varepsilon_{St+s} \xrightarrow{d} \frac{\sigma^2}{S}\int_0^1 \mathbf{W}(r)'\,\mathbf{\Psi}(1)'\,C_j^* d\mathbf{W}(r) \tag{A.22}$$

$$= \frac{\sigma^2 a_j}{2}\int_0^1 \mathbf{W}^{\dagger}(r)'\,C_j d\mathbf{W}^{\dagger}(r) + \frac{\sigma^2 b_j}{2}\int_0^1 \mathbf{W}^{\dagger}(r)'\,C_j^* d\mathbf{W}^{\dagger}(r).$$

Turning to the second term in the numerators of (A.15) and (A.16), using the results established in (A.2)-(A.5) we have that

$$\sum_{t=1}^{N}\sum_{s=1-S}^{0} x_{0,St+s}\left(e^k_{St+s} - \varepsilon_{St+s}\right) = \psi(1)\sum_{t=1}^{N}\sum_{s=1-S}^{0}\left(\sum_{j=1}^{(St+s)-1}\varepsilon_{St+s-j}\right)\sum_{i=k+1}^{\infty}\pi_{k,i}\varepsilon_{St+s-i}$$

and that

$$\sum_{t=1}^{N}\sum_{s=1-S}^{0} x_{S/2,St+s}\left(e^k_{St+s} - \varepsilon_{St+s}\right) = \psi(-1)\sum_{t=1}^{N}\sum_{s=1-S}^{0}\left(\sum_{j=1}^{(St+s)-1}(-1)^j\,\varepsilon_{St+s-j}\right)\sum_{i=k+1}^{\infty}\pi_{k,i}\varepsilon_{St+s-i}$$

and that for $j = 1, ..., S^*$,

$$\begin{aligned}\sum_{t=1}^{N}\sum_{s=1-S}^{0} x_{j,St+s}\left(e^k_{St+s} - \varepsilon_{St+s}\right) &= b_j\sum_{t=1}^{N}\sum_{s=1-S}^{0}\left[\cos(\omega_j[s])\left(\sum_{h=1}^{S}\cos(\omega_j[h-S])\left[\sum_{g=1}^{t-1}\varepsilon_{Sg-(s-h)}\right]\right)\right.\\ &\quad\left.+\sin(\omega_j[s])\left(\sum_{h=1}^{S}\sin(\omega_j[h-S])\left[\sum_{g=1}^{t-1}\varepsilon_{Sg-(s-h)}\right]\right)\right]\sum_{i=k+1}^{\infty}\pi_{k,i}\varepsilon_{St+s-i}\\ &\quad+a_j\sum_{t=1}^{N}\sum_{s=1-S}^{0}\left[\sin(\omega_j[s])\left(\sum_{h=1}^{S}\cos(\omega_j[h-S])\left[\sum_{g=1}^{t-1}\varepsilon_{Sg-(s-h)}\right]\right)\right.\\ &\quad\left.-\cos(\omega_j[s])\left(\sum_{h=1}^{S}\sin(\omega_j[h-S])\left[\sum_{g=1}^{t-1}\varepsilon_{Sg-(s-h)}\right]\right)\right]\sum_{i=k+1}^{\infty}\pi_{k,i}\varepsilon_{St+s-i}\end{aligned}$$

$$\begin{aligned}\sum_{t=1}^{N}\sum_{s=1-S}^{0} x^*_{j,St+s}\left(e^k_{St+s} - \varepsilon_{St+s}\right) &= b_j\sum_{t=1}^{N}\sum_{s=1-S}^{0}\left[\sin(\omega_j[s])\left(\sum_{h=1}^{S}\cos(\omega_j[h-S])\left[\sum_{g=1}^{t-1}\varepsilon_{Sg-(s-h)}\right]\right)\right.\\ &\quad\left.-\cos(\omega_j[s])\left(\sum_{h=1}^{S}\sin(\omega_j[h-S])\left[\sum_{g=1}^{t-1}\varepsilon_{Sg-(s-h)}\right]\right)\right]\sum_{i=k+1}^{\infty}\pi_{k,i}\varepsilon_{St+s-i}\\ &\quad-a_j\sum_{t=1}^{N}\sum_{s=1-S}^{0}\left[\cos(\omega_j[s])\left(\sum_{h=1}^{S}\cos(\omega_j[h-S])\left[\sum_{g=1}^{t-1}\varepsilon_{Sg-(s-h)}\right]\right)\right.\\ &\quad\left.+\sin(\omega_j[s])\left(\sum_{h=1}^{S}\sin(\omega_j[h-S])\left[\sum_{g=1}^{t-1}\varepsilon_{Sg-(s-h)}\right]\right)\right]\sum_{i=k+1}^{\infty}\pi_{k,i}\varepsilon_{St+s-i}\end{aligned}$$

where, in each case, as in Lemma 2 of Berk (1974), we have made use of the substitution $e^k_{St+s} - \varepsilon_{St+s} = \sum_{i=k+1}^{\infty} d_i u_{St+s-i} = \sum_{i=k+1}^{\infty} \pi_{k,i}\varepsilon_{St+s-i}$, where $\sum_{i=k+1}^{\infty} \pi^2_{k,i} \leq \sum_{i=k+1}^{\infty} d_i^2 = o(k^{-2\tau})$, and where τ is as defined in Assumption A.2. It is then straightforward to establish, paralleling the proof of part (a) of Lemma 3.1 in Chang and Park (2002, p.441), that $\sum_{t=1}^N \sum_{s=1-S}^0 x_{j,St+s}\left(e^k_{St+s} - \varepsilon_{St+s}\right) = o_p(T)$ for $j = 0, ..., \lfloor S/2 \rfloor$ and that $\sum_{t=1}^N \sum_{s=1-S}^0 x^*_{j,St+s}\left(e^k_{St+s} - \varepsilon_{St+s}\right) = o_p(T)$ for $j = 1, ..., S^*$. As an immediate consequence of these results, we then obtain, under Assumption A.2,

$$T^{-1}\sum_{t=1}^{N}\sum_{s=1-S}^{0} x_{j,St+s}e^k_{St+s} = T^{-1}\sum_{t=1}^{N}\sum_{s=1-S}^{0} x_{j,St+s}\varepsilon_{St+s} + o_p(1), \quad j = 0, ..., \lfloor S/2 \rfloor \tag{A.23}$$

$$T^{-1}\sum_{t=1}^{N}\sum_{s=1-S}^{0} x^*_{j,St+s}e^k_{St+s} = T^{-1}\sum_{t=1}^{N}\sum_{s=1-S}^{0} x^*_{j,St+s}\varepsilon_{St+s} + o_p(1), \quad j = 1, ..., S^*. \tag{A.24}$$

Combining the results in (A.18)-(A.23), and using applications of the CMT, we obtain the following:

$$T\widehat{\pi}_0 \overset{d}{\to} \frac{\int_0^1 \mathbf{W}^*(r)' C_0 d\mathbf{W}^*(r)}{\psi(1)\int_0^1 \mathbf{W}^*(r)' C_0 \mathbf{W}^*(r)\,dr} \tag{A.25}$$

$$T\widehat{\pi}_{S/2} \overset{d}{\to} \frac{\int_0^1 \mathbf{W}^*(r)' C_{S/2} d\mathbf{W}^*(r)}{\psi(-1)\int_0^1 \mathbf{W}^*(r)' C_{S/2} \mathbf{W}^*(r)\,dr} \tag{A.26}$$

and for $j = 1, ..., S^*$,

$$T\widehat{\pi}_j \overset{d}{\to} \frac{b_j\int_0^1 \mathbf{W}^\dagger(r)' C_j d\mathbf{W}^\dagger(r) - a_j\int_0^1 \mathbf{W}^\dagger(r)' C_j^* d\mathbf{W}^\dagger(r)}{\frac{(a_j^2+b_j^2)}{2}\int_0^1 \mathbf{W}^\dagger(r)' C_j \mathbf{W}^\dagger(r)\,dr} \tag{A.27}$$

$$T\widehat{\pi}^*_j \overset{d}{\to} \frac{a_j\int_0^1 \mathbf{W}^\dagger(r)' C_j d\mathbf{W}^\dagger(r) + b_j\int_0^1 \mathbf{W}^\dagger(r)' C_j^* d\mathbf{W}^\dagger(r)}{\frac{(a_j^2+b_j^2)}{2}\int_0^1 \mathbf{W}^\dagger(r)' C_j \mathbf{W}^\dagger(r)\,dr}. \tag{A.28}$$

Next observe that the corresponding t-statistics from (2.9) can be written as

$$t_j = \hat{\sigma}^{-1}T\widehat{\pi}_i \times \sqrt{T^{-2}\sum_{t=1}^{N}\sum_{s=1-S}^{0}(x_{j,St+s})^2} + o_p(1), \quad j = 0, \ldots, \lfloor S/2 \rfloor \tag{A.29}$$

$$t^*_h = \hat{\sigma}^{-1}T\widehat{\pi}^*_h \times \sqrt{T^{-2}\sum_{t=1}^{N}\sum_{s=1-S}^{0}\left(x^*_{h,St+s}\right)^2} + o_p(1), \quad h = 1, \ldots, S^* \tag{A.30}$$

where $\hat{\sigma}^2$ is the usual OLS variance estimator from (2.9); that is, $\hat{\sigma}^2 := T^{-1}\sum_{t=1}^N \sum_{s=1-S}^0 (\hat{e}^k_{St+s})^2 = T^{-1}\left(\mathbf{u}'\mathbf{u} - \mathbf{u}'\mathbf{Z}_k\left(\mathbf{Z}_k'\mathbf{Z}_k\right)^{-1}\mathbf{Z}_k'\mathbf{u}\right)$. It then follows immediately from parts (i) and (ii) of Lemma A.1 that $\hat{\sigma}^2 = T^{-1}\mathbf{u}'\mathbf{u} + o_p(1)$. Then since part (c) of Lemma 3.1 in Chang and Park (2002) also applies here, we obtain the result that $\hat{\sigma}^2 \overset{p}{\to} \sigma^2$. Substituting this result together with those in (A.25)-(A.28) into (A.29)-(A.30) and using applications of the CMT we then, after some simple manipulations, finally obtain the stated results in Proposition 1, where we have defined the independent standard Brownian motions, $B_i(r) := \mathbf{v}_i'\mathbf{W}^*(r)$, $i = 0, S/2$, $B_j(r) := \mathbf{c}_j'\mathbf{W}^\dagger(r)$ and $B^*_j(r) := \mathbf{c}^{*\prime}_j\mathbf{W}^\dagger(r)'$, where $\mathbf{c}_j'$ and $\mathbf{c}^{*\prime}_j$ are the first rows of $\mathbf{v}_j'$ and $\mathbf{v}^{*\prime}_j$, respectively for $j = 1, \ldots, S^*$.

Table 1: Empirical size of quarterly HEGY tests for white noise innovations

	t_0	t_2	F_1	F_{12}	F_{012}	t_0	t_2	F_1	F_{12}	F_{012}
					A: $N = 50$					
	$\ell = 4$					$\ell = 12$				
AIC	0.048	0.046	0.054	0.056	0.060	0.047	0.049	0.054	0.054	0.060
BIC	0.046	0.045	0.052	0.052	0.057	0.046	0.047	0.051	0.050	0.054
MAIC	0.038	0.032	0.039	0.039	0.043	0.029	0.030	0.032	0.034	0.034
MBIC	0.038	0.032	0.040	0.040	0.043	0.030	0.030	0.033	0.034	0.035
SAIC	0.050	0.050	0.055	0.058	0.065	0.066	0.064	0.072	0.080	0.088
SBIC	0.047	0.047	0.053	0.054	0.060	0.052	0.051	0.054	0.060	0.066
SMAIC	0.031	0.029	0.035	0.035	0.036	0.014	0.012	0.010	0.012	0.012
SMBIC	0.031	0.029	0.035	0.034	0.036	0.013	0.011	0.009	0.009	0.011
t-sq(5%)	0.048	0.047	0.054	0.057	0.062	0.052	0.052	0.057	0.060	0.066
t-sq(10%)	0.048	0.048	0.055	0.058	0.064	0.053	0.051	0.051	0.058	0.063
t-bm(5%)	0.048	0.048	0.054	0.056	0.063	0.059	0.059	0.064	0.066	0.073
t-bm(10%)	0.050	0.050	0.055	0.060	0.065	0.066	0.066	0.071	0.077	0.084
t-rt(5%)	0.048	0.049	0.054	0.058	0.064	0.058	0.058	0.065	0.070	0.078
t-rt(10%)	0.051	0.051	0.055	0.058	0.066	0.067	0.067	0.074	0.082	0.093
					B: $N = 100$					
	$\ell = 4$					$\ell = 12$				
AIC	0.050	0.058	0.053	0.057	0.057	0.048	0.056	0.052	0.054	0.056
BIC	0.049	0.056	0.053	0.054	0.056	0.045	0.054	0.051	0.052	0.054
MAIC	0.039	0.046	0.043	0.041	0.046	0.033	0.038	0.038	0.036	0.039
MBIC	0.041	0.047	0.043	0.041	0.046	0.034	0.041	0.040	0.039	0.042
SAIC	0.053	0.061	0.058	0.058	0.060	0.060	0.068	0.066	0.076	0.081
SBIC	0.049	0.055	0.054	0.053	0.054	0.047	0.054	0.051	0.062	0.063
SMAIC	0.033	0.036	0.031	0.032	0.036	0.014	0.016	0.013	0.015	0.014
SMBIC	0.032	0.034	0.031	0.031	0.035	0.014	0.014	0.011	0.014	0.013
t-sq(5%)	0.050	0.058	0.054	0.057	0.059	0.051	0.059	0.060	0.059	0.062
t-sq(10%)	0.050	0.059	0.056	0.058	0.059	0.049	0.058	0.057	0.059	0.062
t-bm(5%)	0.049	0.061	0.055	0.058	0.060	0.055	0.062	0.061	0.063	0.065
t-bm(10%)	0.050	0.061	0.057	0.059	0.061	0.060	0.067	0.063	0.070	0.073
t-rt(5%)	0.052	0.060	0.056	0.055	0.059	0.054	0.061	0.061	0.072	0.073
t-rt(10%)	0.054	0.061	0.059	0.059	0.061	0.061	0.067	0.069	0.079	0.081

Notes: The DGP is (4.1) with $c = 0$ and $u_{4t+s} = \varepsilon_{4t+s} \sim IID\ N(0,1)$, for quarterly data over N years. Lag selection criteria for the HEGY test regression (4.4) are described in Section 4, with maximum lag in all cases $k_{\max} = \lfloor \ell\,(4N/100)^{1/4} \rfloor$ for $\ell = 4$ or $\ell = 12$. Unit root test statistics are t-type tests at the zero and Nyquist frequencies (t_0, t_2, respectively) and joint F-type statistics at the harmonic seasonal frequency (F_1), at all seasonal frequencies (F_{12}) and at the zero and all seasonal frequencies (F_{012}). Results show the proportion of rejections over 5000 replications for a nominal test size of 0.05.

Table 2: Empirical size of quarterly HEGY tests for moving average disturbances

	t_0	t_2	F_1	F_{12}	F_{012}	t_0	t_2	F_1	F_{12}	F_{012}
A: MA(1) and MA(2) Processes										
$N = 50$	$\theta_1 = -0.8$, $\theta_2 = 0$, $\Theta = 0$					$\theta_1 = 0$, $\theta_2 = -0.64$, $\Theta = 0$				
AIC	0.052	0.347	0.056	0.287	0.277	0.055	0.056	0.251	0.235	0.226
BIC	0.053	0.546	0.062	0.467	0.442	0.071	0.072	0.424	0.413	0.418
MAIC	0.024	0.078	0.024	0.061	0.061	0.021	0.025	0.059	0.056	0.052
MBIC	0.025	0.093	0.024	0.067	0.066	0.021	0.024	0.069	0.063	0.059
SAIC	0.068	0.289	0.076	0.253	0.257	0.075	0.077	0.173	0.178	0.176
SBIC	0.063	0.474	0.067	0.408	0.399	0.065	0.069	0.278	0.267	0.269
SMAIC	0.017	0.056	0.015	0.039	0.037	0.012	0.017	0.032	0.028	0.023
SMBIC	0.016	0.060	0.014	0.039	0.037	0.011	0.015	0.034	0.029	0.023
t-sq(5%)	0.052	0.229	0.059	0.197	0.188	0.053	0.057	0.145	0.134	0.134
t-sq(10%)	0.050	0.147	0.057	0.127	0.123	0.051	0.056	0.098	0.095	0.097
t-bm(5%)	0.054	0.316	0.058	0.265	0.252	0.063	0.061	0.190	0.186	0.173
t-bm(10%)	0.055	0.263	0.058	0.223	0.214	0.063	0.063	0.164	0.163	0.155
t-rt(5%)	0.065	0.354	0.074	0.312	0.305	0.070	0.076	0.204	0.200	0.197
t-rt(10%)	0.068	0.281	0.076	0.247	0.249	0.076	0.079	0.170	0.173	0.174
$N = 100$										
AIC	0.048	0.220	0.055	0.161	0.148	0.055	0.057	0.138	0.126	0.121
BIC	0.050	0.385	0.063	0.308	0.286	0.053	0.054	0.261	0.226	0.210
MAIC	0.027	0.062	0.027	0.046	0.045	0.031	0.030	0.046	0.044	0.041
MBIC	0.026	0.075	0.026	0.053	0.051	0.029	0.028	0.054	0.051	0.047
SAIC	0.064	0.180	0.071	0.152	0.152	0.068	0.069	0.107	0.106	0.110
SBIC	0.056	0.327	0.066	0.261	0.248	0.058	0.059	0.168	0.155	0.143
SMAIC	0.021	0.044	0.015	0.024	0.023	0.018	0.015	0.024	0.022	0.022
SMBIC	0.018	0.046	0.012	0.025	0.025	0.017	0.013	0.027	0.023	0.023
t-sq(5%)	0.051	0.144	0.056	0.114	0.107	0.053	0.057	0.088	0.090	0.084
t-sq(10%)	0.051	0.099	0.055	0.079	0.076	0.050	0.055	0.071	0.073	0.068
t-bm(5%)	0.049	0.186	0.055	0.137	0.130	0.057	0.061	0.109	0.109	0.103
t-bm(10%)	0.048	0.153	0.056	0.118	0.111	0.057	0.060	0.101	0.100	0.095
t-rt(5%)	0.059	0.213	0.068	0.175	0.174	0.066	0.066	0.116	0.116	0.116
t-rt(10%)	0.063	0.171	0.071	0.147	0.145	0.067	0.069	0.104	0.102	0.105

Table 2 (continued)

	t_0	t_2	F_1	F_{12}	F_{012}	t_0	t_2	F_1	F_{12}	F_{012}
					B: Annual MA Processes					
$N = 50$	$\theta_1 = 0,\ \theta_2 = 0,\ \Theta = 0.5$					$\theta_1 = 0,\ \theta_2 = 0,\ \Theta = -0.5$				
AIC	0.217	0.224	0.244	0.303	0.348	0.078	0.070	0.076	0.085	0.104
BIC	0.446	0.453	0.533	0.653	0.710	0.082	0.077	0.080	0.089	0.112
$MAIC$	0.051	0.057	0.052	0.072	0.085	0.022	0.014	0.025	0.018	0.019
$MBIC$	0.053	0.060	0.055	0.075	0.088	0.021	0.014	0.032	0.022	0.020
$SAIC$	0.157	0.164	0.171	0.204	0.214	0.078	0.077	0.091	0.099	0.109
$SBIC$	0.189	0.194	0.222	0.257	0.275	0.076	0.074	0.085	0.096	0.109
$SMAIC$	0.030	0.031	0.028	0.030	0.027	0.012	0.007	0.007	0.006	0.007
$SMBIC$	0.026	0.028	0.026	0.027	0.025	0.009	0.006	0.008	0.005	0.006
t-sq(5%)	0.116	0.124	0.121	0.163	0.193	0.064	0.057	0.058	0.068	0.075
t-sq(10%)	0.094	0.101	0.095	0.126	0.150	0.061	0.053	0.058	0.062	0.070
t-bm(5%)	0.171	0.180	0.190	0.229	0.243	0.073	0.066	0.080	0.086	0.097
t-bm(10%)	0.158	0.172	0.173	0.206	0.219	0.080	0.067	0.083	0.089	0.098
t-rt(5%)	0.161	0.170	0.190	0.222	0.234	0.077	0.077	0.091	0.098	0.108
t-rt(10%)	0.155	0.163	0.170	0.201	0.212	0.077	0.078	0.091	0.101	0.110
$N = 100$										
AIC	0.127	0.133	0.132	0.178	0.212	0.054	0.051	0.056	0.059	0.060
BIC	0.193	0.209	0.200	0.289	0.350	0.079	0.078	0.074	0.085	0.099
$MAIC$	0.046	0.056	0.041	0.048	0.055	0.027	0.025	0.027	0.026	0.022
$MBIC$	0.050	0.059	0.042	0.052	0.061	0.021	0.016	0.025	0.020	0.016
$SAIC$	0.100	0.110	0.098	0.114	0.119	0.066	0.067	0.070	0.072	0.073
$SBIC$	0.123	0.130	0.133	0.160	0.172	0.056	0.054	0.056	0.055	0.056
$SMAIC$	0.023	0.025	0.021	0.021	0.017	0.013	0.013	0.007	0.008	0.007
$SMBIC$	0.021	0.024	0.019	0.019	0.015	0.009	0.010	0.005	0.005	0.005
t-sq(5%)	0.080	0.084	0.078	0.097	0.111	0.050	0.049	0.049	0.052	0.047
t-sq(10%)	0.067	0.077	0.067	0.079	0.092	0.049	0.049	0.049	0.052	0.048
t-bm(5%)	0.107	0.111	0.105	0.123	0.129	0.055	0.056	0.058	0.055	0.053
t-bm(10%)	0.102	0.102	0.099	0.115	0.120	0.059	0.057	0.059	0.059	0.057
t-rt(5%)	0.103	0.112	0.102	0.123	0.127	0.063	0.063	0.068	0.068	0.070
t-rt(10%)	0.099	0.107	0.099	0.113	0.116	0.065	0.069	0.072	0.074	0.072

Notes: As for Table 1, except that the DGP has moving average disturbances, with $u_{4t+s} = (1 - \theta_1 L - \theta_2 L^2)(1 - \Theta L^4)\varepsilon_{4t+s}$ and all results are based on $\ell = 12$.

Table 3: Empirical size of quarterly HEGY tests for autoregressive disturbances

	t_0	t_2	F_1	F_{12}	F_{012}	t_0	t_2	F_1	F_{12}	F_{012}
A: Annual AR Processes										
$N = 50$	$\phi = 0,\ \Phi = 0.5$					$\phi = 0,\ \Phi = -0.5$				
AIC	0.047	0.046	0.060	0.056	0.059	0.044	0.047	0.050	0.050	0.058
BIC	0.044	0.044	0.061	0.056	0.059	0.066	0.065	0.071	0.083	0.098
$MAIC$	0.022	0.020	0.035	0.027	0.027	0.029	0.032	0.028	0.031	0.034
$MBIC$	0.019	0.016	0.037	0.027	0.024	0.029	0.033	0.028	0.031	0.034
$SAIC$	0.076	0.075	0.093	0.099	0.106	0.073	0.076	0.084	0.085	0.083
$SBIC$	0.061	0.058	0.070	0.070	0.074	0.057	0.059	0.066	0.069	0.061
$SMAIC$	0.015	0.011	0.009	0.009	0.011	0.015	0.016	0.015	0.016	0.014
$SMBIC$	0.012	0.008	0.008	0.008	0.009	0.013	0.015	0.014	0.015	0.013
t-sq(5%)	0.053	0.048	0.063	0.064	0.067	0.048	0.049	0.053	0.055	0.061
t-sq(10%)	0.055	0.051	0.063	0.067	0.069	0.050	0.051	0.057	0.056	0.064
t-bm(5%)	0.059	0.057	0.074	0.072	0.073	0.067	0.068	0.071	0.075	0.069
t-bm(10%)	0.067	0.063	0.083	0.081	0.082	0.073	0.074	0.080	0.081	0.080
t-rt(5%)	0.070	0.068	0.085	0.090	0.095	0.066	0.067	0.076	0.077	0.073
t-rt(10%)	0.078	0.075	0.095	0.100	0.110	0.076	0.079	0.088	0.087	0.086
$N = 100$										
AIC	0.051	0.050	0.050	0.051	0.052	0.052	0.054	0.052	0.053	0.056
BIC	0.050	0.049	0.048	0.048	0.051	0.052	0.050	0.050	0.051	0.055
$MAIC$	0.036	0.032	0.033	0.034	0.030	0.040	0.037	0.036	0.035	0.039
$MBIC$	0.025	0.020	0.031	0.028	0.023	0.041	0.039	0.038	0.037	0.040
$SAIC$	0.073	0.064	0.076	0.079	0.087	0.069	0.072	0.068	0.074	0.077
$SBIC$	0.057	0.053	0.060	0.060	0.062	0.057	0.057	0.057	0.059	0.061
$SMAIC$	0.014	0.015	0.009	0.013	0.010	0.018	0.017	0.014	0.015	0.015
$SMBIC$	0.010	0.011	0.007	0.010	0.007	0.018	0.017	0.012	0.013	0.014
t-sq(5%)	0.054	0.052	0.053	0.056	0.059	0.052	0.054	0.054	0.057	0.057
t-sq(10%)	0.055	0.053	0.054	0.056	0.059	0.054	0.054	0.055	0.058	0.059
t-bm(5%)	0.058	0.053	0.057	0.060	0.063	0.064	0.060	0.067	0.066	0.065
t-bm(10%)	0.061	0.056	0.064	0.064	0.069	0.068	0.067	0.069	0.070	0.071
t-rt(5%)	0.068	0.061	0.074	0.074	0.082	0.067	0.066	0.064	0.068	0.072
t-rt(10%)	0.073	0.065	0.077	0.081	0.088	0.070	0.074	0.071	0.077	0.080

Table 3 (continued)

	t_0	t_2	F_1	F_{12}	F_{012}	t_0	t_2	F_1	F_{12}	F_{012}
Panel B: AR (1, 4, 5) Processes										
$N = 50$	$\phi = 0.8$, $\Phi = 0.5$					$\phi = -0.8$, $\Phi = -0.5$				
AIC	0.059	0.053	0.071	0.067	0.076	0.047	0.042	0.051	0.054	0.055
BIC	0.058	0.050	0.067	0.065	0.073	0.048	0.041	0.049	0.051	0.056
MAIC	0.045	0.027	0.039	0.037	0.043	0.031	0.021	0.029	0.032	0.032
MBIC	0.042	0.023	0.037	0.033	0.037	0.032	0.020	0.030	0.032	0.032
SAIC	0.088	0.084	0.115	0.118	0.128	0.101	0.070	0.102	0.097	0.092
SBIC	0.079	0.062	0.085	0.084	0.102	0.093	0.054	0.097	0.089	0.087
SMAIC	0.038	0.004	0.007	0.007	0.016	0.022	0.008	0.023	0.019	0.020
SMBIC	0.041	0.003	0.006	0.005	0.014	0.022	0.007	0.023	0.018	0.019
t-sq(5%)	0.064	0.057	0.076	0.077	0.084	0.054	0.044	0.055	0.058	0.061
t-sq(10%)	0.067	0.059	0.076	0.080	0.084	0.053	0.045	0.054	0.060	0.060
t-bm(5%)	0.077	0.057	0.080	0.072	0.084	0.121	0.049	0.129	0.110	0.113
t-bm(10%)	0.075	0.063	0.084	0.075	0.090	0.104	0.053	0.113	0.099	0.096
t-rt(5%)	0.089	0.075	0.106	0.104	0.118	0.093	0.064	0.098	0.090	0.086
t-rt(10%)	0.089	0.084	0.117	0.119	0.128	0.101	0.070	0.102	0.098	0.092
$N = 100$										
AIC	0.056	0.048	0.068	0.065	0.067	0.045	0.049	0.050	0.050	0.049
BIC	0.054	0.047	0.066	0.061	0.063	0.044	0.048	0.048	0.048	0.047
MAIC	0.040	0.035	0.045	0.041	0.044	0.032	0.034	0.033	0.035	0.033
MBIC	0.047	0.030	0.040	0.035	0.043	0.035	0.032	0.033	0.032	0.033
SAIC	0.070	0.073	0.096	0.095	0.099	0.075	0.064	0.074	0.073	0.068
SBIC	0.061	0.053	0.074	0.067	0.073	0.064	0.054	0.060	0.058	0.054
SMAIC	0.029	0.008	0.012	0.012	0.017	0.019	0.012	0.019	0.015	0.017
SMBIC	0.035	0.003	0.007	0.005	0.013	0.023	0.010	0.019	0.014	0.018
t-sq(5%)	0.057	0.052	0.070	0.068	0.071	0.048	0.053	0.053	0.056	0.052
t-sq(10%)	0.057	0.052	0.070	0.067	0.072	0.050	0.056	0.053	0.057	0.053
t-bm(5%)	0.060	0.048	0.070	0.065	0.069	0.064	0.055	0.062	0.060	0.050
t-bm(10%)	0.060	0.050	0.075	0.072	0.072	0.068	0.057	0.063	0.063	0.053
t-rt(5%)	0.069	0.065	0.091	0.085	0.093	0.072	0.062	0.068	0.067	0.062
t-rt(10%)	0.071	0.074	0.097	0.096	0.100	0.078	0.065	0.076	0.076	0.072

Notes: As for Table 1, except that the DGP has autoregressive disturbances with $(1 - \phi L)(1 - \Phi L^4)u_{4t+s} = \varepsilon_{4t+s}$ and all results are based on $\ell = 12$.

Table 4: Size-corrected power of quarterly HEGY tests for white noise innovations

	t_0	t_2	F_1	F_{12}	F_{012}	t_0	t_2	F_1	F_{12}	F_{012}
A: $N = 50$ years										
	$\ell = 4$					$\ell = 12$				
AIC	0.354	0.348	0.572	0.743	0.839	0.354	0.323	0.526	0.705	0.804
BIC	0.366	0.348	0.587	0.755	0.853	0.352	0.333	0.550	0.724	0.827
$MAIC$	0.359	0.360	0.575	0.735	0.821	0.353	0.350	0.538	0.661	0.772
$MBIC$	0.366	0.367	0.585	0.745	0.830	0.357	0.357	0.558	0.689	0.794
$SAIC$	0.363	0.359	0.578	0.737	0.831	0.318	0.312	0.479	0.639	0.741
$SBIC$	0.361	0.353	0.589	0.753	0.841	0.322	0.331	0.527	0.683	0.794
$SMAIC$	0.394	0.394	0.596	0.739	0.817	0.359	0.350	0.548	0.637	0.688
$SMBIC$	0.397	0.392	0.601	0.736	0.818	0.355	0.352	0.542	0.632	0.680
t-sq(5%)	0.354	0.348	0.583	0.728	0.838	0.299	0.301	0.456	0.625	0.728
t-sq(10%)	0.351	0.347	0.563	0.726	0.833	0.268	0.272	0.431	0.565	0.680
t-bm(5%)	0.365	0.356	0.583	0.740	0.832	0.340	0.339	0.523	0.677	0.788
t-bm(10%)	0.364	0.360	0.576	0.730	0.829	0.318	0.326	0.482	0.665	0.771
t-rt(5%)	0.368	0.358	0.585	0.736	0.831	0.329	0.311	0.507	0.670	0.763
t-rt(10%)	0.363	0.356	0.579	0.735	0.826	0.312	0.315	0.478	0.630	0.727
B: $N = 100$ years										
	$\ell = 4$					$\ell = 12$				
AIC	0.315	0.296	0.548	0.721	0.839	0.323	0.297	0.548	0.716	0.829
BIC	0.323	0.299	0.552	0.736	0.839	0.335	0.303	0.554	0.726	0.829
$MAIC$	0.340	0.304	0.555	0.738	0.838	0.327	0.299	0.537	0.720	0.812
$MBIC$	0.340	0.306	0.556	0.751	0.846	0.358	0.310	0.571	0.740	0.840
$SAIC$	0.314	0.308	0.535	0.711	0.828	0.323	0.292	0.499	0.628	0.768
$SBIC$	0.323	0.306	0.553	0.735	0.847	0.322	0.307	0.553	0.678	0.807
$SMAIC$	0.344	0.337	0.606	0.757	0.858	0.341	0.325	0.601	0.690	0.766
$SMBIC$	0.350	0.344	0.623	0.759	0.855	0.335	0.329	0.610	0.687	0.756
t-sq(5%)	0.319	0.293	0.545	0.711	0.829	0.298	0.270	0.467	0.651	0.762
t-sq(10%)	0.313	0.290	0.533	0.709	0.827	0.280	0.254	0.430	0.609	0.708
t-bm(5%)	0.328	0.296	0.548	0.711	0.828	0.322	0.302	0.549	0.705	0.818
t-bm(10%)	0.326	0.291	0.551	0.713	0.830	0.319	0.305	0.527	0.681	0.807
t-rt(5%)	0.321	0.309	0.555	0.716	0.829	0.320	0.302	0.514	0.647	0.773
t-rt(10%)	0.310	0.305	0.530	0.711	0.828	0.311	0.281	0.491	0.627	0.765

Notes: As for Table 1, except that the DGP is (4.1) with $c = 10$.

Table 5: Size-corrected power of quarterly HEGY tests for moving average disturbances

	t_0	t_2	F_1	F_{12}	F_{012}	t_0	t_2	F_1	F_{12}	F_{012}
A: MA(1) and MA(2) Processes										
$N = 50$	$\theta_1 = -0.8$, $\theta_2 = 0$, $\Theta = 0$					$\theta_1 = 0$, $\theta_2 = -0.64$, $\Theta = 0$				
AIC	0.285	0.400	0.481	0.463	0.553	0.285	0.289	0.727	0.783	0.812
BIC	0.267	0.327	0.508	0.514	0.522	0.195	0.181	0.665	0.692	0.774
MAIC	0.240	0.255	0.412	0.462	0.552	0.269	0.250	0.467	0.553	0.642
MBIC	0.235	0.243	0.422	0.457	0.542	0.280	0.257	0.443	0.543	0.617
SAIC	0.266	0.378	0.424	0.525	0.588	0.298	0.282	0.651	0.707	0.751
SBIC	0.245	0.370	0.474	0.490	0.498	0.282	0.272	0.729	0.807	0.840
SMAIC	0.179	0.267	0.391	0.462	0.519	0.322	0.302	0.489	0.621	0.642
SMBIC	0.177	0.240	0.377	0.439	0.502	0.327	0.302	0.467	0.610	0.631
t-sq(5%)	0.274	0.319	0.438	0.467	0.546	0.271	0.256	0.550	0.622	0.708
t-sq(10%)	0.261	0.307	0.392	0.493	0.589	0.251	0.230	0.523	0.616	0.707
t-bm(5%)	0.268	0.440	0.486	0.560	0.611	0.283	0.298	0.710	0.768	0.796
t-bm(10%)	0.269	0.424	0.470	0.599	0.666	0.292	0.290	0.675	0.745	0.784
t-rt(5%)	0.271	0.391	0.437	0.455	0.538	0.295	0.290	0.689	0.727	0.756
t-rt(10%)	0.266	0.383	0.427	0.537	0.602	0.293	0.283	0.651	0.706	0.749
$N = 100$										
AIC	0.292	0.400	0.450	0.638	0.736	0.259	0.256	0.661	0.773	0.842
BIC	0.296	0.377	0.466	0.607	0.698	0.332	0.326	0.696	0.746	0.778
MAIC	0.244	0.304	0.425	0.602	0.703	0.259	0.270	0.560	0.656	0.746
MBIC	0.257	0.318	0.437	0.590	0.698	0.262	0.274	0.580	0.667	0.747
SAIC	0.254	0.375	0.387	0.600	0.711	0.289	0.281	0.593	0.712	0.766
SBIC	0.270	0.386	0.425	0.591	0.677	0.289	0.275	0.654	0.774	0.845
SMAIC	0.237	0.285	0.430	0.560	0.652	0.322	0.335	0.550	0.707	0.744
SMBIC	0.254	0.264	0.443	0.531	0.634	0.331	0.351	0.547	0.688	0.717
t-sq(5%)	0.268	0.332	0.417	0.584	0.710	0.256	0.239	0.590	0.708	0.785
t-sq(10%)	0.255	0.332	0.395	0.610	0.723	0.256	0.238	0.556	0.660	0.762
t-bm(5%)	0.271	0.398	0.437	0.662	0.769	0.288	0.273	0.650	0.780	0.823
t-bm(10%)	0.280	0.388	0.413	0.664	0.763	0.284	0.283	0.629	0.760	0.803
t-rt(5%)	0.266	0.394	0.403	0.610	0.711	0.295	0.277	0.615	0.737	0.793
t-rt(10%)	0.254	0.369	0.381	0.601	0.710	0.292	0.281	0.591	0.710	0.763

Table 5 (continued)

	t_0	t_2	F_1	F_{12}	F_{012}	t_0	t_2	F_1	F_{12}	F_{012}
					B: Annual MA Processes					
$N = 50$	$\theta_1 = 0,\ \theta_2 = 0,\ \Theta = 0.5$					$\theta_1 = 0,\ \theta_2 = 0,\ \Theta = -0.5$				
AIC	0.755	0.738	0.840	0.856	0.855	0.280	0.292	0.447	0.596	0.682
BIC	0.669	0.629	0.882	0.941	0.964	0.283	0.282	0.456	0.589	0.652
MAIC	0.381	0.343	0.568	0.629	0.692	0.110	0.133	0.190	0.265	0.331
MBIC	0.381	0.344	0.569	0.625	0.683	0.119	0.136	0.178	0.248	0.324
SAIC	0.639	0.610	0.772	0.822	0.829	0.284	0.292	0.404	0.496	0.498
SBIC	0.751	0.751	0.827	0.827	0.798	0.328	0.327	0.482	0.593	0.614
SMAIC	0.659	0.645	0.771	0.757	0.733	0.062	0.066	0.105	0.114	0.118
SMBIC	0.656	0.645	0.770	0.756	0.729	0.025	0.027	0.050	0.042	0.037
t-sq(5%)	0.521	0.499	0.664	0.744	0.799	0.219	0.237	0.355	0.468	0.542
t-sq(10%)	0.432	0.401	0.604	0.689	0.756	0.200	0.212	0.306	0.421	0.499
t-bm(5%)	0.721	0.715	0.850	0.870	0.872	0.309	0.339	0.458	0.540	0.571
t-bm(10%)	0.672	0.646	0.810	0.852	0.860	0.285	0.310	0.423	0.496	0.534
t-rt(5%)	0.674	0.664	0.803	0.837	0.842	0.312	0.303	0.435	0.536	0.543
t-rt(10%)	0.627	0.595	0.764	0.816	0.829	0.277	0.285	0.398	0.479	0.494
$N = 100$										
AIC	0.512	0.477	0.773	0.848	0.878	0.275	0.297	0.427	0.562	0.635
BIC	0.656	0.641	0.811	0.886	0.940	0.342	0.346	0.599	0.758	0.836
MAIC	0.343	0.299	0.595	0.724	0.798	0.251	0.254	0.401	0.527	0.642
MBIC	0.347	0.298	0.611	0.727	0.802	0.111	0.116	0.207	0.260	0.333
SAIC	0.554	0.504	0.740	0.797	0.828	0.293	0.289	0.437	0.561	0.630
SBIC	0.656	0.611	0.809	0.858	0.866	0.303	0.313	0.464	0.595	0.659
SMAIC	0.660	0.645	0.771	0.809	0.807	0.272	0.265	0.437	0.547	0.594
SMBIC	0.670	0.645	0.780	0.822	0.814	0.059	0.053	0.115	0.141	0.140
t-sq(5%)	0.408	0.370	0.645	0.757	0.833	0.234	0.247	0.391	0.525	0.659
t-sq(10%)	0.363	0.320	0.607	0.734	0.828	0.232	0.245	0.388	0.523	0.650
t-bm(5%)	0.599	0.575	0.805	0.853	0.874	0.295	0.298	0.465	0.592	0.667
t-bm(10%)	0.545	0.519	0.759	0.827	0.855	0.292	0.301	0.465	0.574	0.659
t-rt(5%)	0.593	0.536	0.760	0.820	0.850	0.292	0.292	0.435	0.573	0.625
t-rt(10%)	0.538	0.500	0.733	0.792	0.824	0.286	0.293	0.428	0.561	0.618

Notes: As for Table 2, except that the DGP is (4.1) with $c = 10$.

Table 6: Size-corrected power of quarterly HEGY tests for autoregressive disturbances

	t_0	t_2	F_1	F_{12}	F_{012}	t_0	t_2	F_1	F_{12}	F_{012}
A: Annual AR Processes										
$N = 50$	$\phi = 0,\ \Phi = 0.5$					$\phi = 0,\ \Phi = -0.5$				
AIC	0.315	0.305	0.501	0.656	0.759	0.204	0.206	0.283	0.416	0.510
BIC	0.502	0.496	0.570	0.581	0.610	0.196	0.192	0.270	0.392	0.469
$MAIC$	0.283	0.270	0.484	0.608	0.697	0.070	0.064	0.107	0.137	0.154
$MBIC$	0.288	0.277	0.496	0.620	0.709	0.039	0.038	0.086	0.087	0.082
$SAIC$	0.433	0.412	0.586	0.679	0.721	0.213	0.214	0.295	0.397	0.485
$SBIC$	0.491	0.467	0.633	0.689	0.757	0.213	0.210	0.319	0.428	0.517
$SMAIC$	0.471	0.483	0.643	0.660	0.678	0.074	0.082	0.119	0.155	0.159
$SMBIC$	0.473	0.474	0.633	0.656	0.670	0.021	0.022	0.044	0.040	0.032
t-sq(5%)	0.288	0.282	0.466	0.609	0.716	0.202	0.202	0.262	0.381	0.491
t-sq(10%)	0.267	0.250	0.414	0.565	0.679	0.187	0.194	0.244	0.357	0.476
t-bm(5%)	0.482	0.467	0.632	0.704	0.745	0.226	0.239	0.324	0.437	0.530
t-bm(10%)	0.438	0.426	0.605	0.678	0.723	0.220	0.229	0.310	0.423	0.520
t-rt(5%)	0.451	0.438	0.609	0.685	0.728	0.216	0.209	0.306	0.407	0.492
t-rt(10%)	0.423	0.405	0.574	0.673	0.724	0.209	0.216	0.290	0.397	0.485
$N = 100$										
AIC	0.281	0.273	0.500	0.676	0.795	0.246	0.246	0.401	0.546	0.684
BIC	0.283	0.301	0.504	0.679	0.798	0.245	0.248	0.412	0.558	0.687
$MAIC$	0.291	0.305	0.518	0.662	0.769	0.257	0.262	0.406	0.538	0.669
$MBIC$	0.302	0.311	0.531	0.682	0.794	0.027	0.029	0.118	0.111	0.105
$SAIC$	0.355	0.347	0.568	0.659	0.753	0.252	0.249	0.372	0.506	0.607
$SBIC$	0.373	0.376	0.600	0.711	0.781	0.245	0.252	0.393	0.538	0.658
$SMAIC$	0.428	0.425	0.649	0.713	0.750	0.262	0.265	0.441	0.522	0.603
$SMBIC$	0.427	0.430	0.662	0.726	0.751	0.066	0.061	0.127	0.142	0.160
t-sq(5%)	0.279	0.283	0.464	0.635	0.764	0.237	0.243	0.377	0.515	0.652
t-sq(10%)	0.262	0.270	0.448	0.604	0.739	0.226	0.230	0.362	0.492	0.625
t-bm(5%)	0.372	0.375	0.566	0.701	0.789	0.254	0.266	0.423	0.560	0.672
t-bm(10%)	0.359	0.368	0.547	0.699	0.788	0.265	0.267	0.408	0.548	0.653
t-rt(5%)	0.351	0.347	0.570	0.677	0.767	0.258	0.260	0.377	0.510	0.618
t-rt(10%)	0.354	0.342	0.552	0.646	0.750	0.246	0.261	0.372	0.505	0.609

Table 6 (continued)

	t_0	t_2	F_1	F_{12}	F_{012}	t_0	t_2	F_1	F_{12}	F_{012}
	Panel B: AR(1, 4, 5) Processes									
$N = 50$	$\phi = 0.8$, $\Phi = 0.5$					$\phi = -0.8$, $\Phi = -0.5$				
AIC	0.303	0.284	0.495	0.662	0.784	0.094	0.112	0.128	0.175	0.224
BIC	0.312	0.272	0.524	0.683	0.782	0.104	0.111	0.127	0.172	0.223
MAIC	0.254	0.272	0.436	0.578	0.677	0.094	0.079	0.090	0.127	0.166
MBIC	0.254	0.279	0.435	0.578	0.680	0.037	0.017	0.030	0.030	0.038
SAIC	0.486	0.273	0.639	0.724	0.737	0.105	0.112	0.133	0.173	0.200
SBIC	0.571	0.267	0.676	0.756	0.712	0.092	0.100	0.125	0.171	0.195
SMAIC	0.440	0.326	0.554	0.650	0.607	0.056	0.147	0.134	0.200	0.180
SMBIC	0.421	0.336	0.538	0.640	0.596	0.034	0.079	0.054	0.079	0.061
t-sq(5%)	0.266	0.276	0.461	0.619	0.747	0.095	0.109	0.127	0.169	0.208
t-sq(10%)	0.258	0.268	0.431	0.578	0.721	0.091	0.109	0.127	0.164	0.199
t-bm(5%)	0.598	0.281	0.687	0.743	0.692	0.105	0.110	0.134	0.181	0.226
t-bm(10%)	0.545	0.290	0.661	0.738	0.709	0.096	0.117	0.145	0.196	0.225
t-rt(5%)	0.519	0.281	0.652	0.741	0.737	0.098	0.107	0.129	0.166	0.197
t-rt(10%)	0.483	0.273	0.636	0.722	0.733	0.108	0.113	0.132	0.176	0.199
$N = 100$										
AIC	0.335	0.276	0.517	0.703	0.836	0.158	0.169	0.215	0.315	0.418
BIC	0.330	0.278	0.525	0.713	0.838	0.158	0.170	0.219	0.324	0.419
MAIC	0.288	0.270	0.491	0.654	0.777	0.150	0.176	0.220	0.338	0.405
MBIC	0.289	0.295	0.510	0.682	0.796	0.169	0.143	0.169	0.255	0.333
SAIC	0.416	0.286	0.615	0.722	0.791	0.161	0.158	0.199	0.301	0.380
SBIC	0.450	0.297	0.653	0.773	0.815	0.157	0.166	0.212	0.330	0.422
SMAIC	0.429	0.333	0.643	0.720	0.722	0.120	0.251	0.255	0.364	0.392
SMBIC	0.408	0.356	0.637	0.725	0.697	0.108	0.251	0.277	0.386	0.357
t-sq(5%)	0.312	0.264	0.491	0.658	0.800	0.158	0.167	0.206	0.298	0.394
t-sq(10%)	0.295	0.255	0.462	0.627	0.770	0.155	0.163	0.200	0.291	0.384
t-bm(5%)	0.452	0.312	0.658	0.773	0.839	0.158	0.181	0.226	0.331	0.406
t-bm(10%)	0.434	0.307	0.652	0.762	0.836	0.157	0.185	0.221	0.337	0.409
t-rt(5%)	0.422	0.297	0.627	0.748	0.792	0.164	0.161	0.206	0.300	0.380
t-rt(10%)	0.408	0.285	0.601	0.721	0.784	0.160	0.159	0.196	0.295	0.377

Notes: As for Table 3, except that the DGP is (4.1) with $c = 10$.