Queen's Economics Department Working Paper No. 1224

Nearly Efficient Likelihood Ratio Tests for Seasonal Unit Roots

Michael Jansson
UC Berkeley and CREATES

Morten Ørregaard Nielsen
Queen's University and CREATES

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

11-2009

Nearly Efficient Likelihood Ratio Tests for Seasonal Unit Roots*

Michael Jansson†
UC Berkeley and CREATES

Morten Ørregaard Nielsen
Queen's University and CREATES

June 26, 2010

Abstract

In an important generalization of zero frequency autoregressive unit root tests, Hylleberg, Engle, Granger & Yoo (1990) developed regression-based tests for unit roots at the seasonal frequencies in quarterly time series. We develop likelihood ratio tests for seasonal unit roots and show that these tests are "nearly efficient" in the sense of Elliott, Rothenberg & Stock (1996), i.e. that their asymptotic local power functions are indistinguishable from the Gaussian power envelope. Nearly efficient testing procedures for seasonal unit roots have been developed, including point optimal tests based on the Neyman-Pearson Lemma as well as regression-based tests, e.g. Rodrigues & Taylor (2007). However, both require the choice of a GLS detrending parameter, which our likelihood ratio tests do not.

*Prepared for a special issue of *Journal of Time Series Econometrics* in honor of Svend Hylleberg. We are grateful to the editors, three anonymous referees, Rob Taylor, and participants at the CREATES conference on Periodicity, Non-stationarity, and Forecasting of Economic and Financial Time Series for comments and discussion, and to the Social Sciences and Humanities Research Council of Canada (SSHRC grant no. 410-2009-0183) and the Center for Research in Econometric Analysis of Time Series (CREATES, funded by the Danish National Research Foundation) for financial support.

†Corresponding author. Address: University of California, Berkeley, Department of Economics, 508-1 Evans Hall #3880, Berkeley, CA 94720. Email: mjansson@econ.berkeley.edu

1 Introduction

Determining the number and locations of unit roots in non-annual economic time series is a problem that has attracted considerable attention over the last couple of decades. In a important generalization of the work of Dickey and Fuller (1979, 1981) and Dickey, Hasza & Fuller (1984), Hylleberg et al. (1990, henceforth HEGY) developed regression-based tests of the subhypotheses comprising the seasonal unit root hypothesis in a quarterly context. Subsequent work has further generalized the HEGY tests in various ways, including to models with seasonal intercepts and/or trends and to non-quarterly seasonal models (e.g., Beaulieu & Miron (1993), Rodrigues & Taylor (2004), and Smith, Taylor & Castro (2009)).

From the point of view of statistical efficiency, the properties of the HEGY tests are analogous to those of their zero frequency counterparts, the Dickey-Fuller tests. In particular, in models without deterministic components the HEGY t-tests are "nearly efficient" in the sense of Elliott et al. (1996, henceforth ERS), i.e. their asymptotic local power functions are indistinguishable from the Gaussian power envelope. However, the HEGY t-tests are asymptotically inefficient in models with intercepts and/or trends. To improve power of seasonal unit root tests, Gregoir (2006) and Rodrigues & Taylor (2007, henceforth RT) have extended the asymptotic power envelopes of ERS to seasonal models and have developed feasible tests that are nearly efficient in seasonal contexts. As do their zero frequency counterparts due to ERS, the nearly efficient tests of Gregoir (2006) and RT involve so-called GLS detrending, implementation of which requires the choice of a vector of "non-centrality" parameters. The purpose of this paper is to propose nearly efficient seasonal unit root tests that enjoy the (aesthethically as well as scientifically) appealing feature that they do not require the choice of such non-centrality parameters.

To do so, we generalize the analysis of Jansson & Nielsen (2009, henceforth JN), who propose nearly efficient likelihood ratio tests of the zero frequency unit root hypothesis, to models appropriate for testing for seasonal unit roots. Specifically, the paper proceeds as follows. Section 2 is concerned with testing for seasonal unit roots in quarterly time series in the simplest possible setting, namely a Gaussian AR(4) model with standard normal innovations and with presample observations assumed to be equal to their expected values. We develop likelihood ratio unit root tests in this model and show that these tests are nearly efficient. Section 3 discusses extensions to models with serially correlated and/or non-Gaussian errors and to tests for seasonal unit roots in non-quarterly time series. Proofs of our results are provided in Section 4.

2 Likelihood Ratio Tests for Seasonal Unit Roots

2.1 No Deterministic Component

Suppose $\{y_t : 1 \leq t \leq T\}$ is an observed univariate quarterly time series generated by the zero-mean Gaussian AR(4) model

$$\rho(L) y_t = \varepsilon_t, \tag{1}$$

where $\rho(L)$ is a lag polynomial of order four, $\varepsilon_t \sim i.i.d.\ \mathcal{N}(0,1)$, and the initial conditions are $y_{-3} = \ldots = y_0 = 0$.[1] Following RT we assume that $\rho(L)$ admits the factorization

$$\rho(L) = (1 - \rho_Z L)(1 + \rho_N L)\left(1 + \rho_A L^2\right), \tag{2}$$

where ρ_Z, ρ_N, and ρ_A are (unknown) parameters.[2]

Under the quarterly unit root hypothesis

$$H_0 : \rho_Z = 1, \rho_N = 1, \rho_A = 1,$$

the polynomial $\rho(L)$ simplifies to $\Delta_4 = 1 - L^4$, implying that $\{y_t\}$ is a quarterly random walk process. Defining $H_0^k : \rho_k = 1$ for $k \in \{Z, N, A\}$, the quarterly unit root hypothesis H_0 can be expressed as

$$H_0 = H_0^Z \cap H_0^N \cap H_0^A.$$

The hypotheses H_0^Z and H_0^N correspond to a unit root at the zero and Nyquist frequencies $\omega = 0$ and $\omega = \pi$, respectively, while H_0^A yields a pair of complex conjugate unit roots at the frequencies $\omega = \pi/2$ (i.e., the annual frequency) and $\omega = 3\pi/2$.

The alternative corresponding to the single frequency unit root null hypothesis H_0^k is given by $H_1^k : \rho_k < 1$ for $k \in \{Z, N, A\}$. However, we consider also the intermediate alternative hypotheses $H_{1,0}^Z : \rho_Z < 1, \rho_N = \rho_A = 1$, $H_{1,0}^N : \rho_N < 1, \rho_Z = \rho_A = 1$,

[1] The initial values assumption can be relaxed to $\max(|y_{-3}|, \ldots, |y_0|) = o_P\left(\sqrt{T}\right)$ without invalidating the asymptotic results reported in Theorem 1 below.

[2] In the notation of RT, we study a model with periodicity $S = 4$ and parameters ρ_Z, ρ_N, and ρ_A given by $\rho_Z = \alpha_0$, $\rho_N = \alpha_2$, and $\rho_A = \alpha_1^2$, respectively. The local-to-unity parameters in Theorems 1 and 2 of this paper are related to those in $(2.5) - (2.6)$ of RT as follows: $c_Z = c_0$, $c_N = c_2$, and $c_A = c_1 + O\left(T^{-1}\right)$.

and $H_{1,0}^{A}:\rho_A<1,\rho_Z=\rho_N=1$, where unit roots are assumed present at the frequencies not being tested.

Specifically, the likelihood ratio test statistic associated with the problem of testing H_0 vs. $H_{1,0}^{Z}:\rho_Z<1,\rho_N=\rho_A=1$ is given by

$$LR_T^Z=\max_{\bar{\rho}_Z\leq 1}L_T\left(\bar{\rho}_Z,1,1\right)-L_T\left(1,1,1\right),$$

where $L_T\left(\rho_Z,\rho_N,\rho_A\right)=-\sum_{t=1}^{T}\left[\left(1-\rho_Z L\right)\left(1+\rho_N L\right)\left(1+\rho_A L^2\right)y_t\right]^2/2$ is the log likelihood function. Developing a likelihood ratio test of H_0^Z under the "as if" assumption that $\rho_N=\rho_A=1$ is analytically convenient because $L_T\left(\cdot,1,1\right)$ is a quadratic function. Moreover, because Remark 3.2 of RT shows that the large sample properties of the point optimal test statistics $L_T\left(1+T^{-1}\bar{c}_Z,1,1\right)-L_T\left(1,1,1\right)$ are invariant with respect to local departures of ρ_N and/or ρ_A from unity (for any $\bar{c}_Z$), it seems plausible that a similar invariance property will be enjoyed by LR_T^Z. Theorem 1 below confirms this conjecture and further shows that the test which rejects for large values of LR_T^Z is a nearly efficient test of H_0^Z vs. $H_1^Z:\rho_Z<1$.

By analogy with LR_T^Z, define

$$LR_T^N=\max_{\bar{\rho}_N\leq 1}L_T\left(1,\bar{\rho}_N,1\right)-L_T\left(1,1,1\right)$$

and

$$LR_T^A=\max_{\bar{\rho}_A\leq 1}L_T\left(1,1,\bar{\rho}_A\right)-L_T\left(1,1,1\right).$$

As defined, LR_T^N is the likelihood ratio test statistic associated with the problem of testing H_0 vs. $H_{1,0}^{N}:\rho_N<1,\rho_Z=\rho_A=1$, but it will be shown below that the test based on LR_T^N is nearly efficient when testing H_0^N vs. $H_1^N:\rho_N<1$. Again, asymptotic invariance of LR_T^N with respect to local departures of ρ_Z and/or ρ_A from unity is expected in light of the invariance result for point optimal test statistics reported in Remark 3.2 of RT. Similarly, it turns out that a nearly efficient test of H_0^A vs. $H_1^A:\rho_A<1$ can be based on LR_T^A, the likelihood ratio test statistic associated with the problem of testing H_0 vs. $H_{1,0}^{A}:\rho_A<1,\rho_Z=\rho_N=1$.

Note that the alternative hypotheses for our likelihood ratio tests are composite, e.g. $\rho_Z<1$ for the zero frequency test. On the other hand, the alternatives for the nearly efficient tests in RT are point alternatives, e.g. $\rho_Z=\bar{\rho}_Z<1$.

To characterize the local-to-unity asymptotic behavior of the likelihood ratio statistics LR_T^Z, LR_T^N, and LR_T^A, we proceed as in JN. For $k\in\{Z,N,A\}$, the likelihood ratio statistic LR_T^k admits a representation of the form

$$LR_T^k = \max_{\bar{c} \leq 0} \left[\bar{c} S_T^k - \frac{1}{2} \bar{c}^2 H_T^k \right], \quad (3)$$

where S_T^k and H_T^k are the score and Hessian, respectively, of the log-likelihood function $L_T(\rho_Z, \rho_N, \rho_A)$ with respect to ρ_k, $k \in \{Z, N, A\}$, evaluated under the null hypothesis, see (13)-(15) in the proof of Theorem 1. The large-sample behavior of the pair $\left(S_T^k, H_T^k\right)$ is well understood from the work of RT (and others). As a consequence, we obtain the following result, in which

$$W_{c_k}^k(r) = \int_0^r \exp\left[c_k(r-s)\right] dW^k(s), \qquad k = Z, N, A,$$

where $W^Z(\cdot)$, $W^N(\cdot)$, and $W^A(\cdot)$ are independent Wiener processes of dimensions 1, 1, and 2, respectively.

Theorem 1 *Suppose $\{y_t\}$ is generated by (1). If $c_Z = T(\rho_Z - 1)$, $c_N = T(\rho_N - 1)$, and $c_A = T(\rho_A - 1)/2$ are held fixed as $T \to \infty$, then the following hold jointly:*

$$LR_T^k \to_d \max_{\bar{c} \leq 0} \Lambda_{c_k}^k(\bar{c}) \text{ for } k = Z, N, A,$$

where

$$\Lambda_{c_k}^k(\bar{c}) = \bar{c} \cdot tr\left[\int_0^1 W_{c_k}^k(r) dW_{c_k}^k(r)'\right] - \frac{1}{2}\bar{c}^2 tr\left[\int_0^1 W_{c_k}^k(r) W_{c_k}^k(r)' dr\right].$$

Theorem 1 implies in particular that the local asymptotic properties of each LR_T^k depends on the local-to-unity parameters (c_Z, c_N, c_A) only through c_k. This result, which is unsurprising in light of Remark 3.2 of RT, provides a (partial) statistical justification for developing tests of each H_0^k under the "as if" assumption that the parameters not under test are equal to unity, as it implies that LR_T^k is asymptotically pivotal under H_0^k. In particular, the test which rejects when LR_T^k exceeds κ has asymptotic null rejection probability given by $\Pr\left[\max_{\bar{c} \leq 0} \Lambda_0^k(\bar{c}) > \kappa\right]$ under the assumptions of Theorem 1. Therefore, if $\alpha \leq \Pr\left[\max_{\bar{c} \leq 0} \Lambda_0^k(\bar{c}) > 0\right]$ then the (asymptotic) size α test based on LR_T^k has a critical value $\kappa_{LR}^k(\alpha)$ defined by the requirement $\Pr\left[\max_{\bar{c} \leq 0} \Lambda_0^k(\bar{c}) > \kappa_{LR}^k(\alpha)\right] = \alpha$.[3]

In addition to being asymptotically pivotal under H_0^k, the statistic LR_T^k enjoys the property that it can be used to perform nearly efficient tests of H_0^k vs. H_1^k. In the

[3]The condition $\alpha \leq \Pr\left[\max_{\bar{c} \leq 0} \Lambda_0^k(\bar{c}) > 0\right]$ is satisfied at conventional significance levels since $\Pr\left[\max_{\bar{c} \leq 0} \Lambda_0^Z(\bar{c}) > 0\right] = \Pr\left[\max_{\bar{c} \leq 0} \Lambda_0^N(\bar{c}) > 0\right] \approx 0.6827$ and $\Pr\left[\max_{\bar{c} \leq 0} \Lambda_0^A(\bar{c}) > 0\right] \approx 0.6322$.

case of $k \in \{Z, N\}$, this optimality result follows from Theorem 3.1 of RT and the discussion following Theorem 1 of JN. Moreover, a variant of the same argument establishes optimality when $k = A$. For completeness, we briefly discuss the $k = A$ case here. In all cases, we can exploit the fact (also used in the proof of Theorem 1) that $\max_{\bar{c} \leq 0} \Lambda^k_{c_k}(\bar{c})$ admits the representation

$$\max_{\bar{c} \leq 0} \Lambda^k_{c_k}(\bar{c}) = \frac{\min\left(tr\left[\int_0^1 W^k_{c_k}(r)\, dW^k_{c_k}(r)'\right], 0\right)^2}{2tr\left[\int_0^1 W^k_{c_k}(r)\, W^k_{c_k}(r)'\, dr\right]}. \tag{4}$$

The representation (4) shows that (for conventional significance levels) the test based on LR^A_T is asymptotically equivalent to the HEGY t-test, which in turn implies that the likelihood ratio test is nearly efficient because it follows from Gregoir (2006, Figure 1) and Theorem 3.1 of RT that the HEGY t-test is nearly efficient in the absence of deterministic terms.

Theorem 1 is mostly of theoretical interest, as the model (1) makes a number of unrealistic simplifying assumptions, including (a) the assumption that deterministics are absent, (b) the assumption that the errors ε_t are $i.i.d.\ \mathcal{N}(0,1)$, and (c) the assumption $y_{-3} = \ldots = y_0 = 0$ made about the most recent presample values. The assumption that deterministics are absent will be relaxed in the next subsection, while Section 3.1 will describe how certain types of serial correlation and/or an unknown error distribution can be accommodated. In assuming $y_{-3} = \ldots = y_0 = 0$, we are following Gregoir (2006) and RT as well as most of the literature on zero frequency unit roots and cointegration, e.g. ERS and Johansen (1995, Chapter 14). As is well understood (e.g., RT), the initial values assumption can be relaxed to $\max(|y_{-3}|, \ldots, |y_0|) = o_P\left(\sqrt{T}\right)$ without invalidating the asymptotic results reported in Theorem 1. Similarly, Theorem 2 below remains valid if the initial values assumption made in (5) is relaxed to $\max(|u_{-3}|, \ldots, |u_0|) = o_P\left(\sqrt{T}\right)$. On the other hand, different distributional results and hence different local power properties will generally be obtained if $\max(|y_{-3}|, \ldots, |y_0|) \neq o_P(\sqrt{T})$ in (1) or $\max(|u_{-3}|, \ldots, |u_0|) \neq o_P(\sqrt{T})$ in (5). This has been shown in the context of zero frequency unit root testing by Elliott (1999), Müller & Elliott (2003), and Harvey, Leybourne & Taylor (2009), among others. We leave for future work the development of seasonal analogues of the results obtained in those papers.

Remark. For specificity we have only considered tests for a unit root at a single frequency. Tests of joint hypotheses, such as H_0, can be based on the sum of the relevant single frequency statistics. It is an open question whether such tests are nearly efficient, but because Remark 3.3 of RT shows that a point optimal test statistic for a hypothesis involving multiple frequencies is asymptotically equivalent to the sum of the relevant single frequency (point optimal) test statistics, it is not inconceivable that this might be the case.

2.2 Deterministics

To explore the extent to which the "near efficiency" results of the previous subsection extend to models with deterministics, we consider a model in which $\{y_t : 1 \leq t \leq T\}$ is generated by the Gaussian AR(4) model

$$y_t = \beta' d_t + u_t, \qquad \rho(L) u_t = \varepsilon_t, \tag{5}$$

where $d_t = 1$ or $d_t = (1, t)'$, β is an unknown parameter, $\rho(L)$ is parameterized as in (2), $\varepsilon_t \sim i.i.d.\ \mathcal{N}(0,1)$, and $u_{-3} = \ldots = u_0 = 0.$[4]

In this case, the log likelihood function $L_T^d(\cdot)$ is conveniently expressed as

$$L_T^d(\rho_Z, \rho_N, \rho_A, \beta) = -\frac{1}{2}\left(Y_\rho - D_\rho \beta\right)' \left(Y_\rho - D_\rho \beta\right),$$

where, setting $y_{-3} = \ldots = y_0 = 0$ and $d_{-3} = \ldots = d_0 = 0$, Y_ρ and D_ρ are matrices with row $t = 1, \ldots, T$ given by $\rho(L) y_t$ and $\rho(L) d_t'$, respectively.

The likelihood ratio test associated with the problem of testing H_0 vs. $H_{1,0}^Z$ rejects for large values of

$$\begin{aligned} LR_T^{Z,d} &= \max_{\bar{\rho}_Z \leq 1, \beta} L_T^d(\bar{\rho}_Z, 1, 1, \beta) - L_T^d(1, 1, 1, \beta) \\ &= \max_{\bar{\rho}_Z \leq 1} \mathscr{L}_T^d(\bar{\rho}_Z, 1, 1) - \mathscr{L}_T^d(1, 1, 1), \end{aligned}$$

where

$$\begin{aligned} \mathscr{L}_T^d(\rho_Z, \rho_N, \rho_A) &= \max_\beta L_T^d(\rho_Z, \rho_N, \rho_A, \beta) \\ &= -\frac{1}{2} Y_\rho' Y_\rho + \frac{1}{2}\left(Y_\rho' D_\rho\right)\left(D_\rho' D_\rho\right)^{-1}\left(D_\rho' Y_\rho\right) \end{aligned}$$

[4]To conserve space we do not consider seasonal frequency intercepts and/or trends. Accommodating such d_t should be conceptually straightforward, but is left for future research.

is the profile log likelihood function obtained by maximizing $L_T^d(\rho_Z, \rho_N, \rho_A, \beta)$ with respect to the nuisance parameter β. Analogously, the likelihood ratio statistics associated with tests of H_0 against $H_{1,0}^N$ and $H_{1,0}^A$ are given by

$$LR_T^{N,d} = \max_{\bar{\rho}_N \leq 1} \mathcal{L}_T^d(1, \bar{\rho}_N, 1) - \mathcal{L}_T^d(1,1,1)$$

and

$$LR_T^{A,d} = \max_{\bar{\rho}_A \leq 1} \mathcal{L}_T^d(1, 1, \bar{\rho}_A) - \mathcal{L}_T^d(1,1,1),$$

respectively.

As in the case of LR_T^k, the large sample behavior of $LR_T^{k,d}$ can be analyzed by proceeding as in JN.

Theorem 2 *Suppose $\{y_t\}$ is generated by (5) and suppose $c_Z = T(\rho_Z - 1)$, $c_N = T(\rho_N - 1)$, and $c_A = T(\rho_A - 1)/2$ are held fixed as $T \to \infty$.*

(a) If $d_t = 1$, then the following hold jointly:

$$LR_T^{k,d} \to_d \max_{\bar{c} \leq 0} \Lambda_{c_k}^k(\bar{c}) \text{ for } k = Z, N, A.$$

(b) If $d_t = (1,t)'$, then the following hold jointly:

$$LR_T^{k,d} \to_d \max_{\bar{c} \leq 0} \Lambda_{c_k}^k(\bar{c}) \text{ for } k = N, A$$

and

$$LR_T^{Z,d} \to_d \max_{\bar{c} \leq 0} \Lambda_{c_Z}^{Z,\tau}(\bar{c}),$$

where

$$\Lambda_{c_Z}^{Z,\tau}(\bar{c}) = \Lambda_{c_Z}^Z(\bar{c}) + \frac{1}{2} \frac{\left[(1-\bar{c}) W_{c_Z}^Z(1) + \bar{c}^2 \int_0^1 r W_{c_Z}^Z(r)\, dr\right]^2}{1 - \bar{c} + \bar{c}^2/3} - \frac{1}{2} W_{c_Z}^Z(1)^2.$$

Table 1: Simulated critical values of the $LR_T^{k,d}$ statistic

T	80%	85%	90%	95%	97.5%	99%	99.5%	99.9%
Panel A: $k \in \{Z, N\}$ without trend or $k = N$ with trend								
100	1.14	1.47	1.93	2.69	3.42	4.34	5.02	6.54
200	0.93	1.24	1.68	2.44	3.19	4.15	4.87	6.50
400	0.83	1.10	1.49	2.21	2.94	3.91	4.64	6.32
1000	0.79	1.02	1.37	2.01	2.69	3.61	4.32	5.99
∞	0.76	0.98	1.31	1.88	2.48	3.29	3.92	5.40
Panel B: $k = A$ with or without trend								
100	0.68	0.90	1.22	1.78	2.35	3.12	3.70	5.03
200	0.69	0.91	1.24	1.81	2.40	3.19	3.79	5.21
400	0.69	0.92	1.25	1.83	2.43	3.23	3.85	5.29
1000	0.70	0.93	1.26	1.84	2.44	3.25	3.86	5.32
∞	0.70	0.93	1.26	1.84	2.45	3.27	3.90	5.38
Panel C: $k = Z$ with trend								
100	2.90	3.29	3.81	4.65	5.44	6.44	7.16	8.74
200	2.73	3.11	3.64	4.51	5.34	6.40	7.18	8.91
400	2.59	2.97	3.48	4.34	5.17	6.25	7.05	8.85
1000	2.51	2.86	3.35	4.17	4.98	6.04	6.83	8.65
∞	2.45	2.79	3.26	4.05	4.82	5.82	6.57	8.30

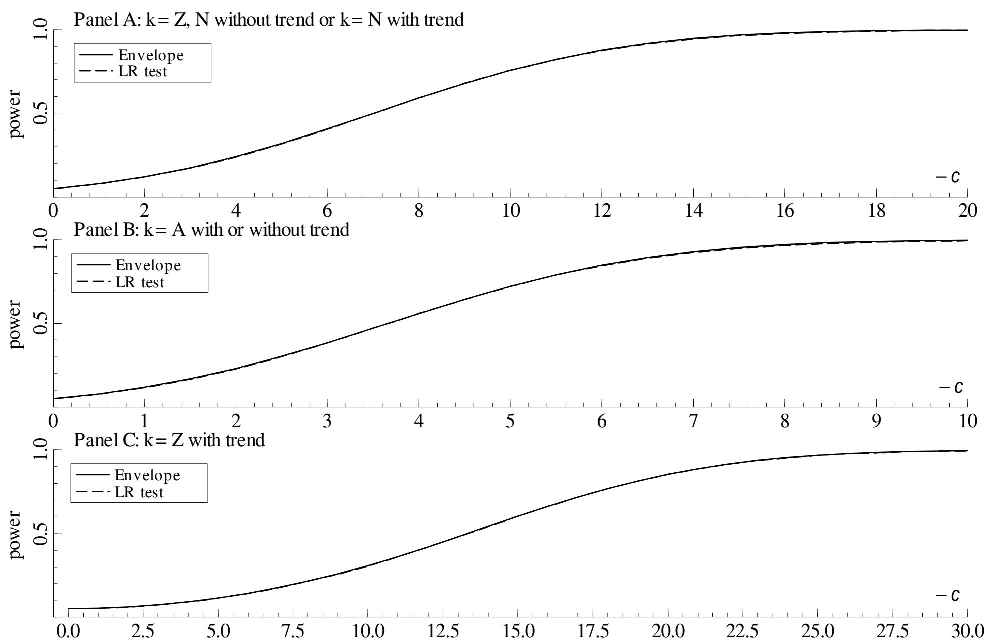
Note: Entries for finite T are simulated quantiles of $LR_T^{k,d}$ with $\varepsilon_t \sim i.i.d.\mathcal{N}(0,1)$. In Panel A it is the $k = Z$ test that is simulated. Entries for $T = \infty$ are simulated quantiles of the corresponding asymptotic distributions, where Wiener processes are approximated by $10,000$ discrete steps with standard Gaussian white noise innovations. All entries are based on ten million Monte Carlo replications.

It follows from Theorem 2 that each $LR_T^{k,d}$ enjoys properties that are qualitatively similar to those enjoyed by LR_T^k in the model without deterministics. Specifically, Theorem 2 implies that each $LR_T^{k,d}$ is asymptotically pivotal under H_0^k. Moreover, Theorem 3.2 of RT and the discussion following Theorem 2 of JN implies that $LR_T^{k,d}$ can be used to perform nearly efficient tests of H_0^k vs. H_1^k.

Simulated critical values $\kappa_{LR}^{k,d}(\alpha)$ associated with $LR_T^{k,d}$ are reported in Table 1.

The profile log likelihood function $\mathscr{L}_T^d(\rho_Z, \rho_N, \rho_A)$ is invariant under transformations of the form $y_t \to y_t + b' d_t$, so that $LR_T^{k,d}$ and any other test statistic that can expressed as a functional of $\mathscr{L}_T^d(\rho_Z, \rho_N, \rho_A)$ shares this invariance property. It therefore makes sense to compare the asymptotic local power properties of the

Figure 1: Power envelope and asymptotic local power of seasonal unit root LR tests



Note: Simulated power envelopes and asymptotic local power functions based on one million Monte Carlo replications, where Wiener processes were approximated by $T = 10,000$ discrete steps with standard Gaussian white noise innovations.

likelihood ratio tests $LR_T^{k,d}$ with the Gaussian power envelopes for invariant tests derived in ERS, Gregoir (2006), and RT.

The asymptotic local power function (with argument $c \leq 0$) of the size α likelihood ratio test is given by $\Pr[\max_{\bar{c}\leq 0} \Lambda_{c_k}^{k}(\bar{c}) > \kappa_{LR}^{k,d}(\alpha)]$ in case of $d_t = 1$ (any k) or $d_t = (1,t)', k = N, A$ and by $\Pr[\max_{\bar{c}\leq 0} \Lambda_{c_Z}^{Z,\tau}(\bar{c}) > k_{LR}^{Z,\tau}(\alpha)]$ in case of $d_t = (1,t)', k = Z$, where $\kappa_{LR}^{k,d}(\alpha)$ satisfies $\Pr[\max_{\bar{c}\leq 0} \Lambda_{0}^{k,d}(\bar{c}) > \kappa_{LR}^{k,d}(\alpha)] = \alpha$ and $\kappa_{LR}^{Z,\tau}(\alpha)$ satisfies $\Pr[\max_{\bar{c}\leq 0} \Lambda_{0}^{Z,\tau}(\bar{c}) > \kappa_{LR}^{Z,\tau}(\alpha)] = \alpha$. Figure 1 plots these functions for $\alpha = 0.05$ in the three cases: $k \in \{Z, N\}$ without trend or $k = N$ with trend (Panel A), $k = A$ with or without trend (Panel B), and $k = Z$ with trend (Panel C). Also plotted in each panel of Figure 1 are the corresponding Gaussian power envelopes, which (for size α tests) are given by $\Pr[\Lambda_{c_k}^{k}(\bar{c}) > \kappa_{\bar{c}}^{k,d}(\alpha)]\big|_{\bar{c}=c_k}$ in case of $d_t = 1$ (any k) or $d_t = (1,t)', k = N, A$ and by $\Pr[\Lambda_{c_Z}^{Z,\tau}(\bar{c}) > \kappa_{\bar{c}}^{Z,\tau}(\alpha)]\big|_{\bar{c}=c_Z}$ in case of $d_t =$

$(1,t)', k=Z$, where $\kappa_{\bar{c}}^{k,d}(\alpha)$ satisfies $\Pr[\Lambda_0^{k,d}(\bar{c}) > \kappa_{\bar{c}}^{k,d}(\alpha)] = \alpha$ and $\kappa_{\bar{c}}^{Z,\tau}(\alpha)$ satisfies $\Pr[\Lambda_0^{Z,\tau}(\bar{c}) > \kappa_{\bar{c}}^{Z,\tau}(\alpha)] = \alpha$.

In each panel of Figure 1, the asymptotic local power functions of the likelihood ratio tests are indistinguishable from the Gaussian power envelopes, so that near optimality claims can be made on the part of the likelihood ratio tests for each case. To avoid cluttering the figure we have not plotted the asymptotic local power functions of the modified point optimal invariant tests and GLS-HEGY tests of RT. However, if plotted, these would also appear indistinguishable from the Gaussian power envelope, see RT (Remark 5.2). In addition, the asymptotic local power functions of the OLS-HEGY tests can be found in Rodrigues & Taylor (2004).

3 Extensions

The results of the previous section can be generalized in a variety of ways. This section briefly discusses two such extensions.

3.1 Serial Correlation and Unknown Error Distribution

One natural extension is to relax the AR(4) specification and the normality assumption on the part of the innovations $\{\varepsilon_t\}$. To that end, suppose $\{y_t : 1 \le t \le T\}$ is generated by the model

$$y_t = \beta' d_t + u_t, \qquad \rho(L)\gamma(L)u_t = \varepsilon_t, \tag{6}$$

where $d_t = 1$ or $d_t = (1,t)'$, β is an unknown parameter, $\rho(L)$ is parameterized as in (2), $\gamma(L) = 1 - \gamma_1 L - \ldots - \gamma_p L^p$ is a lag polynomial of (known, finite) order p satisfying $\min_{|z|\le 1}|\gamma(z)| > 0$, the initial conditions are $u_{-p-3} = \ldots = u_0 = 0$, and the ε_t are $i.i.d.$ errors from a distribution with mean zero and unknown variance σ^2.

In this case, the Gaussian quasi-log likelihood function can be expressed as

$$L_T^d(\rho_Z,\rho_N,\rho_A,\beta;\sigma^2,\gamma) = -\frac{T}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\left(Y_{\rho,\gamma} - D_{\rho,\gamma}\beta\right)'\left(Y_{\rho,\gamma} - D_{\rho,\gamma}\beta\right),$$

where, setting $y_{-p-3} = \ldots = y_0 = 0$ and $d_{-p-3} = \ldots = d_0 = 0$, $Y_{\rho,\gamma}$ and $D_{\rho,\gamma}$ are matrices with row $t = 1,\ldots,T$ given by $\rho(L)\gamma(L)y_t$ and $\rho(L)\gamma(L)d_t'$, respectively. The profile quasi-log likelihood function obtained by profiling out β is given by

$$\mathscr{L}_T^d(\rho_Z,\rho_N,\rho_A;\sigma^2,\gamma) = \max_\beta L_T^d(\rho_Z,\rho_N,\rho_A,\beta;\sigma^2,\gamma)$$

$$= -\frac{T}{2}\log\sigma^2 - \frac{1}{2\sigma^2}Y'_{\rho,\gamma}Y_{\rho,\gamma} + \frac{1}{2\sigma^2}\left(Y'_{\rho,\gamma}D_{\rho,\gamma}\right)\left(D'_{\rho,\gamma}D_{\rho,\gamma}\right)^{-1}\left(D'_{\rho,\gamma}Y_{\rho,\gamma}\right).$$

By analogy with JN, it seems natural to consider likelihood ratio-type test statistics of the form

$$\widehat{LR}_T^{Z,d} = \max_{\bar{\rho}_Z \le 1} \mathscr{L}_T^d\left(\bar{\rho}_Z, 1, 1; \hat{\sigma}_T^2, \hat{\gamma}_T\right) - \mathscr{L}_T^d\left(1, 1, 1; \hat{\sigma}_T^2, \hat{\gamma}_T\right),$$

$$\widehat{LR}_T^{N,d} = \max_{\bar{\rho}_N \le 1} \mathscr{L}_T^d\left(1, \bar{\rho}_Z, 1; \hat{\sigma}_T^2, \hat{\gamma}_T\right) - \mathscr{L}_T^d\left(1, 1, 1; \hat{\sigma}_T^2, \hat{\gamma}_T\right),$$

$$\widehat{LR}_T^{A,d} = \max_{\bar{\rho}_A \le 1} \mathscr{L}_T^d\left(1, 1, \bar{\rho}_A; \hat{\sigma}_T^2, \hat{\gamma}_T\right) - \mathscr{L}_T^d\left(1, 1, 1; \hat{\sigma}_T^2, \hat{\gamma}_T\right),$$

where $\hat{\sigma}_T^2$ and $\hat{\gamma}_T$ are plug-in estimators of σ^2 and $\gamma = (\gamma_1, \ldots, \gamma_p)'$, respectively.

The statistic $\widehat{LR}_T^{k,d}$ is straightforward to compute, requiring only maximization with respect to the scalar parameter $\bar{\rho}_k$. Proceeding as in the proof of Theorem 3 of JN, it should be possible to show that if $\{y_t\}$ is generated by (6), $c_Z = T(\rho_Z - 1)$, $c_N = T(\rho_N - 1)$, and $c_A = T(\rho_A - 1)/2$ are held fixed as $T \to \infty$ and if

$$\left(\hat{\sigma}_T^2, \hat{\gamma}_T\right) \to_P \left(\sigma^2, \gamma\right), \tag{7}$$

then

$$\widehat{LR}_T^{k,d} \to_d \max_{\bar{c} \le 0} \Lambda_{c_k}^k(\bar{c}) \text{ for } k = Z, N, A \tag{8}$$

if $d_t = 1$, while

$$\widehat{LR}_T^{k,d} \to_d \max_{\bar{c} \le 0} \Lambda_{c_k}^k(\bar{c}) \text{ for } k = N, A \tag{9}$$

and

$$\widehat{LR}_T^{Z,d} \to_d \max_{\bar{c} \le 0} \Lambda_{c_Z}^{Z,\tau}(\bar{c}) \tag{10}$$

when $d_t = (1, t)'$.

Remarks. (i) The consistency condition (7) is mild. For instance, it is satisfied by

$$\hat{\sigma}_T^2 = \frac{1}{T-p-4}\sum_{t=p+5}^{T}\left(\Delta_4 y_t - \hat{\eta}_T' Z_t\right)^2, \qquad \hat{\gamma}_T = (0, I_p)\,\hat{\eta}_T,$$

where

$$\hat{\eta}_T = \left(\sum_{t=p+5}^{T} Z_t Z_t'\right)^{-1}\left(\sum_{t=p+5}^{T} Z_t \Delta y_t\right), \qquad Z_t = (1, \Delta_4 y_{t-1}, \ldots, \Delta_4 y_{t-p})'.$$

(ii) The assumption $u_{-p-3} = \ldots = u_0 = 0$ made when deriving the quasi-log likelihood function can be relaxed to $\max\left(\left|u_{-p-3}\right|, \ldots, |u_0|\right) = o_P\left(\sqrt{T}\right)$ without invalidating $(8)-(10)$.

(iii) While the distributional results $(8)-(10)$ remain valid under departures from normality, relaxing the assumption of normality of the error distribution does affect the shapes of the power envelopes. This has been shown in the context of zero frequency unit root testing by Rothenberg & Stock (1997) and Jansson (2008), among others.

To assess the size control of the likelihood ratio tests in finite samples we conduct a small Monte Carlo experiment. For specificity and because the presence of a negative moving average component is known to be problematic in unit root testing, we consider as in RT the DGP

$$\Delta_4 y_t = (1 + \theta L^2)\varepsilon_t, \tag{11}$$

where $y_0 = y_{-1} = y_{-2} = y_{-3} = 0$ and $\varepsilon_t \sim i.i.d.\ \mathcal{N}(0,1)$. For the parameter θ we consider values $\theta \in \{-0.75, -0.50, \ldots, 0.75\}$. When θ is large and positive there is near-cancellation of the unit root at the annual frequency, whereas when θ is large and negative there is near-cancellation of the unit roots at the zero and Nyquist frequencies. We simulate the model with sample sizes $T \in \{100, 200, 400\}$ and conduct two separate experiments where we allow for a constant mean in one experiment and for a linear trend in the other.

In the simulations the likelihood ratio test $\widehat{LR}_T^d$ is compared with the modified point optimal test (denoted P_k^{GLS}) and GLS-HEGY (denoted t_k^{GLS} and F_A^{GLS}) tests of RT, and OLS-HEGY (denoted t_k^{OLS} and F_A^{OLS}) tests of Hylleberg et al. (1990) using one million replications of the model (11). As in RT the lag length for the HEGY tests is chosen by a general-to-specific approach starting with an initial four, six, and eight lags for $T = 100$, $T = 200$, and $T = 400$, respectively, and progressively deleting those which are insignificant at the 5% level. To calculate the long-run variance in the modified point optimal tests we use an autoregressive spectral density estimator as in RT with the lag length chosen by the GLS-HEGY regression, and to calculate the plug-in values for the likelihood ratio test we use the lag length chosen by the OLS-HEGY regression (the lag lengths chosen by the GLS-HEGY and OLS-HEGY regressions are the same in the vast majority of the replications). The results of the simulations are presented in Table 2 for the constant mean case and Table 3 for the linear trend case.

In both the constant mean and linear trend cases the null rejection frequencies are seen to be quite sensitive to θ, especially so when $T = 100$. Overall, $\widehat{LR}_T^d$ compares very favorably to the point optimal and GLS-HEGY tests of RT in terms of size control, especially for the zero frequency test. However, $\widehat{LR}_T^{Z,d}$ and $\widehat{LR}_T^{N,d}$ are quite conservative for positive values of θ. We interpret this evidence as suggesting that the new tests developed in this paper should be viewed as serious contenders to currently employed seasonal unit root tests.

3.2 Non-Quarterly Models

Another natural extension is to consider a model with periodicity $S \neq 4$. Following RT, a natural generalization of (5) is given by the Gaussian AR(S) model

$$y_t = \beta' d_t + u_t, \qquad \rho(L) u_t = \varepsilon_t, \tag{12}$$

where $d_t = 1$ or $d_t = (1,t)'$, β is an unknown parameter, $u_{1-S} = \ldots = u_0 = 0$, $\varepsilon_t \sim i.i.d.\ \mathcal{N}(0,1)$, and $\rho(L)$ is parameterized as

$$\rho(L) = (1-\rho_Z L)(1+\rho_N L) \prod_{k=1}^{\lfloor (S-1)/2 \rfloor} \left(1 - 2\rho_k \cos\omega_k L + \rho_k^2 L^2\right) \qquad (S \text{ even}),$$

$$\rho(L) = (1-\rho_Z L) \prod_{k=1}^{\lfloor (S-1)/2 \rfloor} \left(1 - 2\rho_k \cos\omega_k L + \rho_k^2 L^2\right) \qquad (S \text{ odd}),$$

where $\omega_k = 2\pi k/S$ for $k = 1, \ldots, \lfloor (S-1)/2 \rfloor$.

Table 2: Simulated size of seasonal unit root tests, constant mean case

		Zero frequency				Nyquist frequency				Annual frequency			
θ	T	$\widehat{LR}_T^{Z,d}$	P_Z^{GLS}	t_Z^{GLS}	t_Z^{OLS}	$\widehat{LR}_T^{N,d}$	P_N^{GLS}	t_N^{GLS}	t_N^{OLS}	$\widehat{LR}_T^{A,d}$	P_A^{GLS}	F_A^{GLS}	F_A^{OLS}
−0.75	100	0.3415	0.4617	0.7085	0.4960	0.2950	0.3102	0.4584	0.4675	0.0620	0.0054	0.0821	0.0832
	200	0.1838	0.2920	0.4830	0.3409	0.1782	0.2207	0.3124	0.3367	0.0396	0.0278	0.0607	0.0626
	400	0.1201	0.1835	0.2997	0.2215	0.1179	0.1572	0.2220	0.2389	0.0300	0.0235	0.0577	0.0589
−0.50	100	0.1064	0.1984	0.3394	0.1576	0.0796	0.1128	0.1734	0.1799	0.0515	0.0228	0.0750	0.0773
	200	0.0478	0.1286	0.2044	0.0943	0.0481	0.0890	0.1135	0.1293	0.0384	0.0357	0.0608	0.0621
	400	0.0375	0.0878	0.1252	0.0600	0.0385	0.0698	0.0835	0.0959	0.0322	0.0356	0.0555	0.0573
−0.25	100	0.1027	0.1360	0.2415	0.0880	0.0495	0.0762	0.1168	0.1198	0.0603	0.0228	0.0794	0.0810
	200	0.0444	0.0938	0.1429	0.0530	0.0372	0.0644	0.0766	0.0906	0.0468	0.0393	0.0599	0.0632
	400	0.0284	0.0715	0.0965	0.0379	0.0301	0.0570	0.0638	0.0729	0.0399	0.0425	0.0526	0.0546
0.00	100	0.0444	0.0814	0.1508	0.0363	0.0185	0.0408	0.0637	0.0615	0.0588	0.0285	0.0642	0.0671
	200	0.0434	0.0707	0.1099	0.0309	0.0252	0.0470	0.0557	0.0608	0.0544	0.0374	0.0564	0.0590
	400	0.0434	0.0592	0.0803	0.0271	0.0320	0.0471	0.0525	0.0574	0.0507	0.0385	0.0526	0.0548
0.25	100	0.0127	0.1033	0.1536	0.0434	0.0117	0.0513	0.0638	0.0610	0.0754	0.0581	0.0790	0.0808
	200	0.0166	0.0865	0.1198	0.0407	0.0150	0.0565	0.0600	0.0641	0.0482	0.0521	0.0580	0.0600
	400	0.0234	0.0685	0.0878	0.0355	0.0218	0.0538	0.0568	0.0600	0.0416	0.0483	0.0526	0.0549
0.50	100	0.0075	0.1124	0.1707	0.0548	0.0062	0.0536	0.0697	0.0667	0.1167	0.1103	0.1124	0.1222
	200	0.0119	0.0889	0.1145	0.0392	0.0106	0.0569	0.0570	0.0615	0.0690	0.0869	0.0741	0.0802
	400	0.0164	0.0635	0.0770	0.0285	0.0159	0.0493	0.0496	0.0543	0.0483	0.0658	0.0578	0.0618
0.75	100	0.0094	0.0530	0.1347	0.0415	0.0077	0.0252	0.0568	0.0555	0.4211	0.4173	0.4175	0.4484
	200	0.0097	0.0961	0.1359	0.0489	0.0084	0.0606	0.0678	0.0714	0.2394	0.2804	0.2420	0.2659
	400	0.0127	0.0492	0.0652	0.0212	0.0117	0.0380	0.0420	0.0463	0.1530	0.1938	0.1539	0.1820

Note: The table presents simulated null rejection frequencies (i.e., simulated size) for the likelihood ratio test $\widehat{LR}_T^{k,d}$, modified point optimal test (P_k^{GLS}) and GLS-HEGY (t_k^{GLS} and F_A^{GLS}) tests of RT, and OLS-HEGY (t_k^{OLS} and F_A^{OLS}) tests of Hylleberg et al. (1990) using one million replications of the model (11) allowing for a constant mean but no trend. All three tests use lag length chosen by a general-to-specific procedure.

Table 3: Simulated size of seasonal unit root tests, linear trend case

		Zero frequency				Nyquist frequency				Annual frequency			
θ	T	$\widehat{LR}_T^{Z,d}$	P_Z^{GLS}	t_Z^{GLS}	t_Z^{OLS}	$\widehat{LR}_T^{N,d}$	P_N^{GLS}	t_N^{GLS}	t_N^{OLS}	$\widehat{LR}_T^{A,d}$	P_A^{GLS}	F_A^{GLS}	F_A^{OLS}
−0.75	100	0.3000	0.5644	0.7208	0.5536	0.3387	0.3360	0.5063	0.5001	0.0618	0.0024	0.0808	0.0903
	200	0.0857	0.3932	0.5746	0.4585	0.1917	0.2342	0.3391	0.3491	0.0368	0.0151	0.0613	0.0660
	400	0.0359	0.2344	0.4008	0.3355	0.1121	0.1606	0.2313	0.2382	0.0255	0.0181	0.0569	0.0615
−0.50	100	0.1255	0.2656	0.3504	0.2374	0.0965	0.1196	0.2007	0.2021	0.0474	0.0136	0.0766	0.0851
	200	0.0114	0.1752	0.2275	0.1500	0.0429	0.0900	0.1219	0.1283	0.0302	0.0281	0.0600	0.0649
	400	0.0060	0.1120	0.1404	0.0925	0.0302	0.0703	0.0865	0.0910	0.0232	0.0314	0.0552	0.0600
−0.25	100	0.1265	0.1687	0.2356	0.1393	0.0502	0.0710	0.1254	0.1260	0.0557	0.0161	0.0788	0.0873
	200	0.0485	0.1223	0.1413	0.0851	0.0328	0.0627	0.0794	0.0851	0.0366	0.0337	0.0602	0.0661
	400	0.0045	0.0890	0.0958	0.0546	0.0211	0.0559	0.0643	0.0668	0.0281	0.0399	0.0521	0.0567
0.00	100	0.0399	0.0959	0.1114	0.0534	0.0173	0.0361	0.0634	0.0624	0.0553	0.0239	0.0632	0.0720
	200	0.0360	0.0860	0.0934	0.0460	0.0237	0.0442	0.0553	0.0569	0.0511	0.0341	0.0561	0.0611
	400	0.0331	0.0711	0.0743	0.0382	0.0304	0.0459	0.0523	0.0536	0.0483	0.0372	0.0524	0.0571
0.25	100	0.0037	0.1590	0.1167	0.0638	0.0094	0.0467	0.0658	0.0645	0.0646	0.0484	0.0752	0.0831
	200	0.0013	0.1238	0.1056	0.0594	0.0104	0.0543	0.0611	0.0619	0.0362	0.0475	0.0571	0.0615
	400	0.0016	0.0902	0.0836	0.0492	0.0150	0.0531	0.0576	0.0578	0.0299	0.0471	0.0521	0.0569
0.50	100	0.0003	0.1855	0.1438	0.0886	0.0041	0.0505	0.0755	0.0737	0.0982	0.0970	0.1062	0.1185
	200	0.0005	0.1437	0.1025	0.0597	0.0070	0.0569	0.0602	0.0610	0.0538	0.0826	0.0732	0.0797
	400	0.0007	0.0882	0.0693	0.0389	0.0103	0.0490	0.0507	0.0516	0.0347	0.0645	0.0573	0.0626
0.75	100	0.0002	0.0840	0.1115	0.0708	0.0051	0.0233	0.0622	0.0615	0.3851	0.4006	0.4128	0.4422
	200	0.0001	0.1686	0.1227	0.0721	0.0049	0.0595	0.0706	0.0709	0.1891	0.2568	0.2256	0.2410
	400	0.0004	0.0656	0.0546	0.0287	0.0073	0.0381	0.0433	0.0438	0.1215	0.1928	0.1533	0.1712

Note: The table presents simulated null rejection frequencies (i.e., simulated size) for the likelihood ratio test $\widehat{LR}_T^{k,d}$, modified point optimal test (P_k^{GLS}) and GLS-HEGY (t_k^{GLS} and F_A^{GLS}) tests of RT, and OLS-HEGY (t_k^{OLS} and F_A^{OLS}) tests of Hylleberg et al. (1990) using one million replications of the model (11) allowing for a linear trend. All three tests use lag length chosen by a general-to-specific procedure.

In perfect analogy with the quarterly case, the profile log likelihood function implied by the model (12) can be expressed as

$$-\frac{1}{2}Y_\rho'Y_\rho+\frac{1}{2}\left(Y_\rho'D_\rho\right)\left(D_\rho'D_\rho\right)^{-1}\left(D_\rho'Y_\rho\right),$$

where, setting $y_{1-S}=\ldots=y_0=0$ and $d_{1-S}=\ldots=d_0=0$, Y_ρ and D_ρ are matrices with row $t=1,\ldots,T$ given by $\rho\left(L\right)y_t$ and $\rho\left(L\right)d_t'$, respectively. Tests of individual unit root hypotheses can be based on the natural counterparts of the $LR_T^{k,d}$ statistics considered in the quarterly case and Theorem 2 should generalize in a natural way to the model (12).[5] Specifically, the results for test statistics associated with ρ_Z and ρ_N should coincide with those for $LR_T^{Z,d}$ and $LR_T^{N,d}$ in the quarterly case, while the test statistics associated with ρ_k, $k=1,\ldots,\lfloor\left(S-1\right)/2\rfloor$, should exhibit the same large sample behavior as $LR_T^{A,d}$ does in the quarterly case.

4 Proofs

4.1 Proof of Theorem 1

Let

$$S_T^Z=\frac{1}{T}\sum_{t=1}^{T}y_{t-1}^Z\Delta_4y_t,\qquad H_T^Z=\frac{1}{T^2}\sum_{t=1}^{T}\left(y_{t-1}^Z\right)^2,\tag{13}$$

$$S_T^N=\frac{1}{T}\sum_{t=1}^{T}y_{t-1}^N\Delta_4y_t,\qquad H_T^N=\frac{1}{T^2}\sum_{t=1}^{T}\left(y_{t-1}^N\right)^2,\tag{14}$$

and

$$S_T^A=\frac{1}{T/2}\sum_{t=1}^{T}y_{t-2}^A\Delta_4y_t,\qquad H_T^A=\frac{1}{\left(T/2\right)^2}\sum_{t=1}^{T}\left(y_{t-2}^A\right)^2,\tag{15}$$

with the definitions $y_t^Z=\left(1+L\right)\left(1+L^2\right)y_t$, $y_t^N=-\left(1-L\right)\left(1+L^2\right)y_t$, and $y_t^A=-\left(1-L\right)\left(1+L\right)y_t$.

The validity of (3) follows from the fact that the log likelihood function $L_T\left(\cdot\right)$ admits the expansions

[5] The statistics derived in the current environment are similar to the $LR_T^{k,d}$ statistics in the sense that they can be expressed as maximizers of rational polynomial functions, so they should be amenable to asymptotic analysis using a slight modification of the proof of Theorem 2.

$$L_T(\bar{\rho}_Z,1,1)=L_T(1,1,1)+T(\bar{\rho}_Z-1)S_T^Z-\frac{1}{2}[T(\bar{\rho}_Z-1)]^2H_T^Z,$$

$$L_T(1,\bar{\rho}_N,1)=L_T(1,1,1)+T(\bar{\rho}_N-1)S_T^N-\frac{1}{2}[T(\bar{\rho}_N-1)]^2H_T^N,$$

$$L_T(1,1,\bar{\rho}_A)=L_T(1,1,1)+\frac{T}{2}(\bar{\rho}_A-1)S_T^A-\frac{1}{2}\left[\frac{T}{2}(\bar{\rho}_A-1)\right]^2H_T^A.$$

Under the assumptions of Theorem 1, the following hold jointly (e.g., RT):

$$\left(S_T^k,H_T^k\right)\to_d\left(\mathscr{S}_{c_k}^k,\mathscr{H}_{c_k}^k\right),\qquad k=Z,N,A, \tag{16}$$

where

$$\mathscr{S}_{c_k}^k=tr\left[\int_0^1 W_{c_k}^k(r)\,dW_{c_k}^k(r)'\right],$$

$$\mathscr{H}_{c_k}^k=tr\left[\int_0^1 W_{c_k}^k(r)\,W_{c_k}^k(r)'\,dr\right].$$

Theorem 1 follows from (3), (16), and the continuous mapping theorem (CMT) because

$$LR_T^k=\max\nolimits_{\bar{c}\le 0}\left[\bar{c}S_T^k-\frac{1}{2}\bar{c}^2H_T^k\right]$$

$$=\frac{\min\left(S_T^k,0\right)^2}{2H_T^k}\to_d\frac{\min\left(\mathscr{S}_{c_k}^k,0\right)^2}{2\mathscr{H}_{c_k}^k}=\max\nolimits_{\bar{c}\le 0}\Lambda_{c_k}^k(\bar{c}),$$

where the second and third equalities use simple facts about quadratic functions.

4.2 Proof of Theorem 2

Because $\mathscr{L}_T^d(\cdot)$ is invariant under transformations of the form $y_t\to y_t+b'd_t$, we can assume without loss of generality that $\beta=0$. The proofs of parts (a) and (b) are very similar, the latter being slightly more involved, so to conserve space we omit the details for part (a). Likewise, the proofs for $k=N$ and $k=A$ are very similar, so to conserve space we omit the details for $k=A$.

Accordingly, suppose $k \in \{Z,N\}$ and $d_t = (1,t)'$. Let y_t^k be as in the proof of Theorem 1 and define $\tilde{d}_{Tt}^Z = (1+L)\left(1+L^2\right)\tilde{d}_{Tt}$ and $\tilde{d}_{Tt}^N = -(1-L)\left(1+L^2\right)\tilde{d}_{Tt}$, where $\tilde{d}_{Tt} = \frac{1}{4}diag(1,1/\sqrt{T})d_t$. The linear trend likelihood ratio statistic can be written as $LR_T^{k,d} = \max_{\bar{c}\leq 0} F\left(\bar{c}, X_T^k\right)$, where

$$X_T^k = \left(S_T^k, H_T^k, A_T^k, B_T^k\right),$$

$$A_T^k = \left[A_T^k(0), A_T^k(1), A_T^k(2)\right],$$

$$B_T^k = \left[B_T^k(0), B_T^k(1), B_T^k(2)\right],$$

for

$$A_T^k(0) = \sum_{t=1}^{T} \Delta_4 \tilde{d}_{Tt} \Delta_4 y_t,$$

$$A_T^k(1) = \frac{1}{T}\sum_{t=1}^{T} (\Delta_4 \tilde{d}_{Tt} y_{t-1}^k + \tilde{d}_{T,t-1}^k \Delta_4 y_t),$$

$$A_T^k(2) = \frac{1}{T^2}\sum_{t=1}^{T} \tilde{d}_{T,t-1}^k y_{t-1}^k,$$

$$B_T^k(0) = \sum_{t=1}^{T} \Delta_4 \tilde{d}_{Tt} \Delta_4 \tilde{d}_{Tt}',$$

$$B_T^k(1) = \frac{1}{T}\sum_{t=1}^{T} (\Delta_4 \tilde{d}_{Tt} \tilde{d}_{T,t-1}^{k\prime} + \tilde{d}_{T,t-1}^k \Delta_4 \tilde{d}_{T,t}'),$$

$$B_T^k(2) = \frac{1}{T^2}\sum_{t=1}^{T} \tilde{d}_{T,t-1}^k \tilde{d}_{T,t-1}^{k\prime},$$

and

$$F(\bar{c},x) = \bar{c}s - \frac{1}{2}\bar{c}^2 h + \frac{1}{2}N(\bar{c},a)' D(\bar{c},b)^{-1} N(\bar{c},a) - \frac{1}{2}N(0,a)' D(0,b)^{-1} N(0,a)$$

with

$$N(\bar{c},a) = N[\bar{c},a(0),a(1),a(2)] = a(0) - \bar{c}a(1) + \bar{c}^2 a(2),$$

$$D(\bar{c},b) = D[\bar{c},b(0),b(1),b(2)] = b(0) - \bar{c}b(1) + \bar{c}^2 b(2).$$

It follows from standard results (e.g., RT) that

$$X_T^k \to_d \mathscr{X}_{c_k}^k = \left(\mathscr{S}_{c_k}^k, \mathscr{H}_{c_k}^k, \mathscr{A}_{c_k}^k, \mathscr{B}^k\right), \qquad k = Z,N,$$

under the assumptions of Theorem 2, where

$$\mathscr{A}_{c_Z}^Z = \left[\begin{pmatrix} \mathscr{Y} \\ W_{c_Z}^Z(1) \end{pmatrix}, \begin{pmatrix} 0 \\ W_{c_Z}^Z(1) \end{pmatrix}, \begin{pmatrix} 0 \\ \int_0^1 r W_{c_Z}^Z(r)\,dr \end{pmatrix}\right],$$

$$\mathscr{B}^Z = \left[\begin{pmatrix} 1/4 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1/3 \end{pmatrix}\right],$$

$$\mathscr{A}_{c_N}^N = \left[\begin{pmatrix} \mathscr{Y} \\ W_{c_N}^N(1) \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}\right],$$

$$\mathscr{B}^N = \left[\begin{pmatrix} 1/4 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\right],$$

with $\mathscr{Y}$ being a random variable independent of $\left[W^Z(\cdot), W^N(\cdot)\right]$ and distributed as $(\varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \varepsilon_4)/4$.

The result now follows as in the proof of Theorem 2 of JN.

References

Beaulieu, J. J. & Miron, J. A. (1993), 'Seasonal unit roots in aggregate U.S. data', *Journal of Econometrics* **55**, 305–328.

Dickey, D. A. & Fuller, W. A. (1979), 'Distribution of the estimators for autoregressive time series with a unit root', *Journal of the American Statistical Association* **74**, 427–431.

Dickey, D. A. & Fuller, W. A. (1981), 'Likelihood ratio statistics for autoregressive time series with a unit root', *Econometrica* **49**, 1057–1072.

Dickey, D. A., Hasza, D. P. & Fuller, W. A. (1984), 'Testing for unit roots in seasonal time series', *Journal of the American Statistical Association* **79**, 355–367.

Elliott, G. (1999), 'Efficient tests for a unit root when the initial observation is drawn from its unconditional distribution', *International Economic Review* **40**, 767–783.

Elliott, G., Rothenberg, T. J. & Stock, J. H. (1996), 'Efficient tests for an autoregressive unit root', *Econometrica* **64**, 813–836.

Gregoir, S. (2006), 'Efficient tests for the presence of a pair of complex conjugate unit roots in real time series', *Journal of Econometrics* **130**, 45–100.

Harvey, D. I., Leybourne, S. J. & Taylor, A. M. R. (2009), 'Unit root testing in practice: dealing with uncertainty over the trend and initial condition', *Econometric Theory* **25**, 587–636.

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

Jansson, M. (2008), 'Semiparametric power envelopes for tests of the unit root hypothesis', *Econometrica* **76**, 1103–1142.

Jansson, M. & Nielsen, M. Ø. (2009), 'Nearly efficient likelihood ratio tests of the unit root hypothesis'. QED working paper 1213, Queen's University.

Johansen, S. (1995), *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. New York: Oxford University Press.

Müller, U. K. & Elliott, G. (2003), 'Tests for unit root and the initial condition', *Econometrica* **71**, 1269–1286.

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

Rodrigues, P. M. M. & Taylor, A. M. R. (2007), ‘Efficient tests of the seasonal unit root hypothesis’, *Journal of Econometrics* **141**, 548–573.

Rothenberg, T. J. & Stock, J. H. (1997), ‘Inference in a nearly integrated autoregressive model with nonnormal innovations’, *Journal of Econometrics* **80**, 269–286.

Smith, R. J., Taylor, A. M. R. & Castro, T. D. B. (2009), ‘Regression-based seasonal unit root tests’, *Econometric Theory* **25**, 527–560.