SCHOOL OF ECONOMICS AND MANAGEMENT
FACULTY OF SOCIAL SCIENCES
AARHUS UNIVERSITY

CREATES
Center for Research in Econometric Analysis of Time Series

CREATES Research Paper 2009-55

Efficient Likelihood Ratio Tests for Seasonal Unit Roots

Michael Jansson and Morten Ørregaard Nielsen

School of Economics and Management
Aarhus University
Bartholins Allé 10, Building 1322, DK-8000 Aarhus C
Denmark

Nearly Efficient Likelihood Ratio Tests for Seasonal Unit Roots*

Michael Jansson†
UC Berkeley and CREATES

Morten Ørregaard Nielsen
Queen's University and CREATES

November 24, 2009

Abstract. In an important generalization of zero frequency autoregressive unit root tests, Hylleberg, Engle, Granger, and 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, and Stock (1996), i.e. that their local asymptotic power functions are indistinguishable from the Gaussian power envelope. Currently available nearly efficient testing procedures for seasonal unit roots are regression-based and require the choice of a GLS detrending parameter, which our likelihood ratio tests do not.

KEYWORDS: Likelihood Ratio Test, Seasonal Unit Root Hypothesis
JEL Codes: C12, C22

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), Hylleberg, Engle, Granger, and Yoo (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., Smith, Taylor, and 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

*Prepared for a special issue of *Journal of Time Series Econometrics* in honor of Svend Hylleberg. We are grateful to 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

particular, in models without deterministic components the HEGY t-tests are "nearly efficient" in the sense of Elliott, Rothenberg, and Stock (1996, henceforth ERS), i.e. their local asymptotic 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, Rodrigues and Taylor (2007, henceforth RT) have extended the asymptotic power envelopes of ERS and Gregoir (2006) 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 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) appealing feature that they do not require the choice of such non-centrality parameters.

To do so, we generalize the analysis of Jansson and 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

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

[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.

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 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

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

where $L_T(\rho_Z, \rho_N, \rho_A) = -\sum_{t=1}^T [(1 - \rho_Z L)(1 + \rho_N L)(1 + \rho_A L^2) y_t]^2 / 2$ 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, as $L_T(\cdot, 1, 1)$ is a quadratic function. Moreover, remark 3.2 of RT suggests that the large sample properties of LR_T^Z should be invariant with respect to local departures of ρ_N and/or ρ_A from unity. 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

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

and

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

[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 $c_A = c_1 + O(T^{-1})$.

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, the asymptotic invariance of LR_T^N with respect to local departures of ρ_Z and/or ρ_A from unity is to expected in light of 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. $H_{1,0}^A : \rho_A < 1, \rho_Z = \rho_N = 1$.

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

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

where 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 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 $\max_{\bar{c}\leq 0} \Lambda_{c_k}^k(\bar{c})$ admits the representation

$$\max_{\bar{c}\leq 0} \Lambda_{c_k}^k(\bar{c}) = \frac{\min\left(tr\left[\int_0^1 W_{c_k}^k(r)\, dW_{c_k}^k(r)'\right], 0\right)^2}{2tr\left[\int_0^1 W_{c_k}^k(r)\, W_{c_k}^k(r)'\, dr\right]}. \quad (4)$$

The representation (4) shows that (for conventional significance levels) the test based on LR_T^A 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.

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 remark 3.3 of RT suggests 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

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

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

[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 $\Pr\left[\max_{\bar{c}\leq 0} \Lambda_0^A(\bar{c}) > 0\right] \approx 0.6322$.

[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.

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

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

$$LR_T^{Z,d} = \max\nolimits_{\bar{\rho}_Z \leq 1, \beta} L_T^d(\bar{\rho}_Z, 1, 1, \beta) - L_T^d(1, 1, 1, \beta)$$

$$= \max\nolimits_{\bar{\rho}_Z \leq 1} \mathcal{L}_T^d(\bar{\rho}_Z, 1, 1) - \mathcal{L}_T^d(1, 1, 1),$$

where

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

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

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

and

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

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:*

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

(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

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

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.

TABLE 1 ABOUT HERE

The profile log likelihood function $\mathcal{L}_T^d(\rho_Z, \rho_N, \rho_A)$ 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 $\mathcal{L}_T^d(\rho_Z, \rho_N, \rho_A)$ shares this invariance property. It therefore makes sense to compare the local asymptotic power properties of the likelihood ratio tests $LR_T^{k,d}$ with the Gaussian power envelopes for invariant tests derived in ERS, Gregoir (2006), and RT.

FIGURE 1 ABOUT HERE

The local asymptotic 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 $\Pr[\Lambda_0^{Z,\tau}(\bar{c}) > \kappa_{\bar{c}}^{Z,\tau}(\alpha)] = \alpha$. The local asymptotic power functions of the likelihood ratio tests are indistinguishable from the Gaussian power envelopes in each case, so that near optimality claims can be made on the part of the likelihood ratio tests.

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

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

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

$$\mathcal{L}_T^d\left(\rho_Z,\rho_N,\rho_A;\sigma^2,\gamma\right)$$

$$=\max_\beta L_T^d\left(\rho_Z,\rho_N,\rho_A,\beta;\sigma^2,\gamma\right)$$

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

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

$$\widehat{LR}_T^{Z,d}=\max_{\bar\rho_Z\le 1}\mathcal{L}_T^d\left(\bar\rho_Z,1,1;\hat\sigma_T^2,\hat\gamma_T\right)-\mathcal{L}_T^d\left(1,1,1;\hat\sigma_T^2,\hat\gamma_T\right),$$

$$\widehat{LR}_T^{N,d}=\max_{\bar\rho_N\le 1}\mathcal{L}_T^d\left(1,\bar\rho_Z,1;\hat\sigma_T^2,\hat\gamma_T\right)-\mathcal{L}_T^d\left(1,1,1;\hat\sigma_T^2,\hat\gamma_T\right),$$

$$\widehat{LR}_T^{A,d}=\max_{\bar\rho_A\le 1}\mathcal{L}_T^d\left(1,1,\bar\rho_A;\hat\sigma_T^2,\hat\gamma_T\right)-\mathcal{L}_T^d\left(1,1,1;\hat\sigma_T^2,\hat\gamma_T\right),$$

where $\hat\sigma_T^2$ and $\hat\gamma_T$ are plug-in estimators of σ^2 and $\gamma=\left(\gamma_1,\ldots,\gamma_p\right)'$, 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), $c_Z=T\left(\rho_Z-1\right)$, $c_N=T\left(\rho_N-1\right)$, and $c_A=T\left(\rho_A-1\right)/2$ 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

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

if $d_t=1$, while

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

and

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

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 $\max\left(|u_{-p-3}|, \ldots, |u_0|\right) = o_P\left(\sqrt{T}\right)$ without invalidating $(8) - (10)$.

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

$$y_t = \beta' d_t + u_t, \qquad \rho(L)\, u_t = \varepsilon_t, \tag{11}$$

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

In perfect analogy with the quarterly case, the profile log likelihood function implied by the model (11) 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 $\rho(L)\,y_t$ and $\rho(L)\,d_t'$, 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 (11).[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, $k=1,\ldots,\lfloor(S-1)/2\rfloor$, 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

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

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

and

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

where $y_t^Z=(1+L)(1+L^2)\,y_t$, $y_t^N=-(1-L)(1+L^2)\,y_t$, and $y_t^A=-(1-L)(1+L)\,y_t$.

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

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

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

[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.

and

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

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

$$\left(S_T^k, H_T^k\right) \to_d \left(\mathcal{S}_{c_k}^k, \mathcal{H}_{c_k}^k\right), \qquad k = Z, N, A, \tag{12}$$

where

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

Theorem 1 follows from (3), (12), and the continuous mapping theorem (CMT) because

$$LR_T^k = \max_{\bar{c}\leq 0}\left[\bar{c}S_T^k - \frac{1}{2}\bar{c}^2 H_T^k\right] = \frac{\min\left(S_T^k, 0\right)^2}{2H_T^k} \to_d \frac{\min\left(\mathcal{S}_{c_k}^k, 0\right)^2}{2\mathcal{H}_{c_k}^k} = \max_{\bar{c}\leq 0}\Lambda_{c_k}^k(\bar{c}),$$

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

4.2. Proof of Theorem 2. Because $\mathcal{L}_T^d(\cdot)$ 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 suppose $d_t = (1,t)'$. Let y_t^k be as in the proof of Theorem 1 and define $\tilde{d}_{Tt}^Z = (1+L)(1+L^2)\tilde{d}_{Tt}$ and $\tilde{d}_{Tt}^N = -(1-L)(1+L^2)\tilde{d}_{Tt}$, 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

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

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

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

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, under the assumptions of Theorem 2,

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

where

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

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

$$\mathcal{A}^N_{c_N} = \left[\begin{pmatrix} \mathcal{Y} \\ W^N_{c_N}(1) \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right],$$

$$\mathcal{B}^N = \left[\begin{pmatrix} K & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \right],$$

$\mathcal{Y} \sim (\varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \varepsilon_4)/4$ is a random variable independent of $\left[W^Z(\cdot), W^N(\cdot)\right]$, and $K = 1/4$ is a positive constant.

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

References

Dickey, D. A., and W. A. Fuller (1979): "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association*, 74, 427–431.

——— (1981): "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root," *Econometrica*, 49, 1057–1072.

Elliott, G., T. J. Rothenberg, and J. H. Stock (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.

Hylleberg, S., R. F. Engle, C. W. J. Granger, and B. S. Yoo (1990): "Seasonal Integration and Cointegration," *Journal of Econometrics*, 44, 215–238.

Jansson, M., and M. Ø. Nielsen (2009): "Nearly Efficient Likelihood Ratio Tests of the Unit Root Hypothesis," QED working paper 1213, Queen's University.

Rodrigues, P. M. M., and A. M. R. Taylor (2007): "Efficient Tests of the Seasonal Unit Root Hypothesis," *Journal of Econometrics*, 141, 548–573.

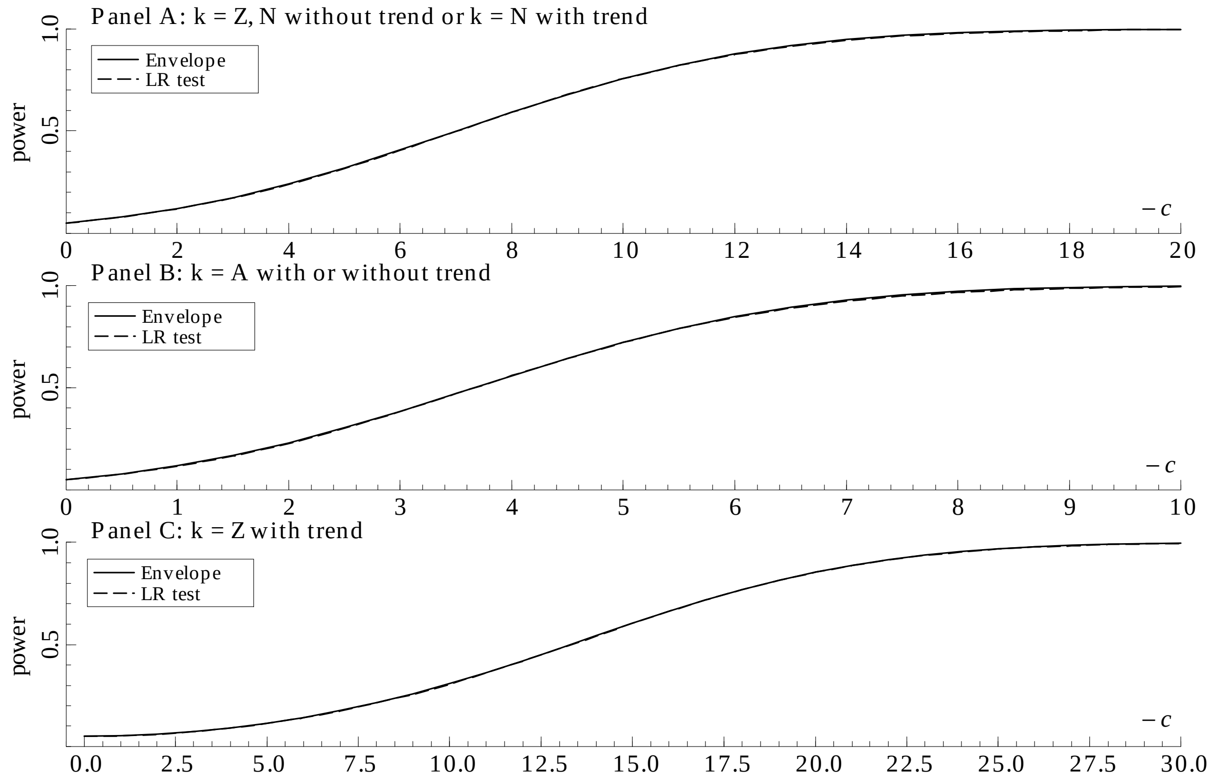
Smith, R. J., A. M. R. Taylor, and T. D. B. Castro (2009): "Regression-Based Seasonal Unit Root Tests," *Econometric Theory*, 25, 527–560.

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	1.1378	1.4706	1.9299	2.6846	3.4110	4.3365	5.0151	6.5243
250	0.8910	1.1824	1.6127	2.3611	3.1074	4.0786	4.8000	6.4382
500	0.8164	1.0681	1.4491	2.1459	2.8709	3.8379	4.5631	6.2526
1000	0.7878	1.0231	1.3726	2.0119	2.6884	3.6152	4.3229	5.9702
∞	0.7612	0.9824	1.3068	1.8831	2.4820	3.2909	3.9180	5.4025
Panel B: $k = A$ with or without trend								
100	0.6781	0.9006	1.2202	1.7806	2.3533	3.1132	3.6949	5.0436
250	0.6901	0.9161	1.2435	1.8200	2.4090	3.2034	3.8084	5.2228
500	0.6946	0.9229	1.2527	1.8333	2.4296	3.2338	3.8523	5.3037
1000	0.6977	0.9257	1.2560	1.8397	2.4393	3.2456	3.8653	5.3322
∞	0.6998	0.9284	1.2604	1.8458	2.4495	3.2670	3.8966	5.3806
Panel C: $k = Z$ with trend								
100	2.9038	3.2898	3.8094	4.6485	5.4402	6.4342	7.1564	8.7342
250	2.6775	3.0598	3.5851	4.4521	5.2893	6.3596	7.1488	8.9060
500	2.5671	2.9327	3.4410	4.2938	5.1225	6.1984	7.0026	8.8210
1000	2.5078	2.8604	3.3510	4.1748	4.9841	6.0401	6.8267	8.6359
∞	2.4541	2.7946	3.2650	4.0512	4.8223	5.8230	6.5795	8.3009

Note: Entries for finite T are simulated quantiles of $LR_T^{k,d}$ with $\varepsilon_t \sim i.i.d.\mathcal{N}(0,1)$, $t = 1, \ldots, T$. 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.

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.

Research Papers 2009

CREATES
Center for Research in Econometric Analysis of Time Series

2009-42: Ole Eiler Barndorff-Nielsen and Robert Stelzer: The multivariate supOU stochastic volatility model

2009-43: Lasse Bork, Hans Dewachter and Romain Houssa: Identification of Macroeconomic Factors in Large Panels

2009-44: Dennis Kristensen: Semiparametric Modelling and Estimation: A Selective Overview

2009-45: Kim Christensen, Silja Kinnebrock and Mark Podolskij: Pre-averaging estimators of the ex-post covariance matrix

2009-46: Matias D. Cattaneo, Richard K. Crump and Michael Jansson: Robust Data-Driven Inference for Density-Weighted Average Derivatives

2009-47: Mark Podolskij and Mathias Vetter: Understanding limit theorems for semimartingales: a short survey

2009-48: Isabel Casas and Irene Gijbels: Unstable volatility functions: the break preserving local linear estimator

2009-49: Torben G. Andersen and Viktor Todorov: Realized Volatility and Multipower Variation

2009-50: Robinson Kruse, Michael Frömmel, Lukas Menkhoff and Philipp Sibbertsen: What do we know about real exchange rate non-linearities?

2009-51: Tue Gørgens, Christopher L. Skeels and Allan H. Würtz: Efficient Estimation of Non-Linear Dynamic Panel Data Models with Application to Smooth Transition Models

2009-52: Torben G. Andersen, Dobrislav Dobrev and Ernst Schaumburg: Jump-Robust Volatility Estimation using Nearest Neighbor Truncation

2009-53: Florian Heinen, Philipp Sibbertsen and Robinson Kruse: Forecasting long memory time series under a break in persistence

2009-54: Tue Gørgens and Allan Würtz: Testing a parametric function against a nonparametric alternative in IV and GMM settings

2009-55: Michael Jansson and Morten Ørregaard Nielsen: Nearly Efficient Likelihood Ratio Tests for Seasonal Unit Roots