The Japanese Economic Review
The Journal of the Japanese Economic Association

The Japanese Economic Review doi: 10.1111/jere.12185
Vol. 70, No. 2, June 2019

IMPROVING THE PERFORMANCE OF A LONG-RUN VARIANCE RATIO TEST FOR A UNIT ROOT

By HUGO FERRER-PÉREZ†, MARÍA-ISABEL AYUDA‡ and ANTONIO AZNAR‡

†Center for Agro-Food Economy and Development
‡University of Zaragoza

Cai and Shintani (2006, *Econometric Theory*, 22, 347–372) considered the impact of introducing an inconsistent long-run variance estimator when constructing a class of kernel-based ratio tests for testing non-stationarity in the series. They found that the quotient of two estimators with different rates of convergence under the null and the alternative hypotheses may lead to a test having an interesting size and power trade-off. This paper develops modified versions of this test, presents new asymptotic results and tabulates critical values. The finite sample performance is explored through Monte Carlo simulations. The results show that the modifications proposed lead to more powerful unit root tests.
JEL Classification Numbers: C12, C22.

1. Introduction

Over the past four decades, the literature on the importance of testing for unit roots in a given economic time series has greatly increased and, in particular, research dedicated to alleviating the frequent concern that many unit root tests exhibit low power and size properties. See the surveys of Haldrup and Jansson (2006) and Patterson (2011), among others.

Cai and Shintani (2006) propose an alternative approach, based on the generalisation of the von Neumann ratio, that consists of constructing a long-run variance ratio test statistic that exploits different rates of convergence of kernel-based estimators under both the null and alternative hypotheses. They constructed four ratio tests that differ in the pair consistency–inconsistency of their components and derive their respective asymptotic behaviour. Through a Monte Carlo simulation study, they assessed the finite sample properties of the tests and concluded that the ratio test that combines a consistent estimator of the long-run variance in the numerator and an inconsistent estimator under the stationary alternative hypothesis in the denominator exhibits (in their words) a good size and reasonable power. This test is the CI test.

Despite this, the CI test has received less attention than it probably deserves. To resolve this, we propose modified CI tests that are more powerful in the presence of autocorrelated errors while achieving size even closer to the nominal level.

To this end, we adopt a twofold strategy. First, we apply two detrending procedures to remove the deterministic component of the series: ordinary least squares (OLS)[1] and local generalised least squares (GLS) by Elliott *et al.* (1996). Second, we replace, in the

[1] As pointed out by Cai and Shintani, the results obtained with their particular demeaning method coincide with those that would have been obtained with OLS-demeaned data. However, this equivalence does not hold in the model with a constant and a linear time trend. In this case, OLS-detrended data produces novel results that have not been derived previously.

numerator, the kernel-based estimator with the autoregressive spectral density estimator for the long-run variance, while maintaining the inconsistency of the kernel estimator in the denominator of the quotient.[2]

The rest of the article is organised as follows. In Section 2 we describe the model used and introduce the original CI test. In Section 3 we propose our modified tests, derive their asymptotic behaviour and tabulate new critical values of the tests. We conduct a comparison simulation study to analyse the finite sample properties in Section 4; Section 5 concludes.

2. Model and long-run variance ratio test statistics

We assume that the data are generated according to the following process (for t = 1,. . ., T):

$$y_t = d_t + v_t, v_t = \rho v_{t-1} + u_t, \tag{1}$$

where the deterministic component is represented by $d_t = \delta' z_t$ and z_t is the unknown deterministic vector. As usual, we focus on the two leading cases: the case in which the model includes an intercept term, $z_t = 1$, denoted with the superscript μ and the case in which the model includes a linear time trend as well, $z_t = (1, t)'$, represented by the superscript τ. The initial value[3] is $v_1 = 0$ and u_t is the error term defined as a zero-mean sequence satisfying $T^{-1/2} \sum_{t=1}^{\lfloor Ts \rfloor} u_t \xrightarrow{d} \omega W(s)$. The long-run variance of u_t is represented by ω^2 and the long-run variance of y_t under the alternative hypothesis is $\omega_y^2 = \omega^2(1-\rho)^{-2}$.

In this paper, we focus on discriminating between the null of a unit root in the series (ρ = 1) and the stationary alternative (ρ < 1).

Cai and Shintani (2006) considered the following non-parametric kernel estimator (Newey and West, 1987) for the long-run variance of any process x_t:

$$\hat{\omega}^2(x_t, K) = \sum_{j=-(K-1)}^{K-1} (1 - |j/K|) T^{-1} \sum_{t=|j|+1}^{T} x_t x_{t-j}, \tag{2}$$

where K is the bandwidth parameter and the Bartlett kernel is used to ensure non-negative estimates,[4] and construct the CI test as follows:

[2] In this paper we do not pursue the variant of replacing the denominator with an inconsistent autoregressive spectral density estimator because the null limiting distribution for the resulting test depends on nuisance parameters, and, hence, its analytical treatment is not as direct as in a non-parametric framework. We will leave this interesting question for future research.

[3] The initial condition assumption has no effects under the null of a unit root but we are well aware of the consequences of assuming a different initial condition under the alternative hypothesis as discussed in Elliott (1999), Elliott and Müller (2006), Harvey and Leybourne (2006) and Harvey *et al.* (2009), among others. This topic has received much attention lately, especially in the design of unit root tests that are robust to the initial condition, but further insights are needed to assess the effect of introducing an inconsistent long-run variance estimator into the behaviour of the long-run variance ratio test. This is beyond the scope of the present paper but it is under current investigation by the authors.

[4] In addition, the Bartlett kernel will ensure the highest power function when the bandwidth parameter is set to the sample size, as demonstrated in Kiefer and Vogelsang (2002).

$$\mathrm{CI}^i = T^2 \frac{\hat{\omega}^2(\Delta y_t^i, K)}{\hat{\omega}^2(y_t^i, T)} \tag{3}$$

for $i = \mu,\tau$, where $\Delta y_t^\mu = \Delta y_t$ is the series in first differences, $\Delta y_t^\tau = \Delta y_t - \overline{\Delta y}, y_t^\mu = y_t - \bar{y}, \bar{y} = T^{-1}\sum_{t=1}^T y_t, y_t^\tau = \check{y}_t - \bar{\check{y}}, \check{y}_t = \sum_{j=1}^t (\Delta y_j - \overline{\Delta y})$ and $\bar{\check{y}} = T^{-1}\sum_{t=1}^T \check{y}_t$.

3. Improving the performance of the CI test

As demonstrated by Cai and Shintani, the CI test offers a good size and power trade-off but it is oversized when a negative moving-average root, and undersized when a positive autoregressive root, underlies the error term in (1) and hence, properties could be improved. Following their suggestions, we attempt to enhance the properties of the test in two directions. One direction is to apply alternative detrending procedures and the second direction is to replace, in the numerator of the ratio test, the kernel-based estimator with the autoregressive spectral density estimator for the long-run variance originally proposed in Berk (1974), while maintaining the inconsistency of the kernel estimator in the denominator of the test. To the best of our knowledge, there is no previous work covering this, so our paper is an attempt to fill this gap in the literature as well as to provide more powerful tests while exploiting the fact that different rates of convergence of different classes of estimators for the long-run variance may lead to tests that can be reliable for testing for unit roots.

3.1 Ordinary least squares and generalised least squares detrending

An appropriate treatment of the deterministic part of the series is crucial to improve the behaviour of unit root tests. Contrary to what is considered in the seminal work, here we apply OLS and local GLS detrending procedures. We define the OLS-filtered series as $\hat{v}_t = y_t - \hat{\delta}' z_t$, where $\hat{\delta}$ is the OLS estimator of δ. We also consider the local GLS procedure that consists of transforming the variables of the model as $y^{\bar{c}} = (y_1, (1-\rho_{\bar{c}}L)y_t)'$ and $z^{\bar{c}} = (z_1, (1-\rho_{\bar{c}}L)z_t)'$ for $t = 2,\ldots,T$ and $\rho_{\bar{c}} = 1 + \bar{c}/T, \bar{c} < 0$.[5] Then, $\tilde{\delta}$ is obtained from regressing $y^{\bar{c}}$ on $z^{\bar{c}}$ by OLS. We define the GLS-filtered series as $\tilde{v}_t = y_t - \tilde{\delta}' z_t$. According to this, we can write

$$\mathrm{CI}_{\mathrm{m}}^i = T^2 \frac{\hat{\omega}^2(\Delta \hat{x}_t^i, K)}{\hat{\omega}^2(\hat{x}_t^i, T)} \tag{4}$$

for m = OLS, GLS and $i = \mu,\tau$. In addition, $\hat{x}_t$ defines a generic filtered series where $\hat{x}_t \equiv \hat{v}_t$ for the OLS case and $\hat{x}_t \equiv \tilde{v}_t$ for the GLS case. We derive the theorem (proved in the appendix) on the null limiting distribution of both modified tests.[6]

---

[5] As suggested by Elliott *et al.* (1996), the parameter $\bar{c}$ is chosen so that the asymptotic local power function of the test is tangent to the power envelope at 50% power, selecting $\bar{c} = -7.0$ for the mean case and $\bar{c} = -13.5$ for the trend case.

[6] Under the near-integrated alternative hypothesis, normally represented as $\mathrm{H}_c : \rho_c = 1 + cT^{-1}$ where $c < 0$ denotes the deviation from the null hypothesis ($c = 0$), the asymptotic distributions are similar to those of Theorem 1 but replacing the Wiener process $W(r)$ with the Ornstein–Uhlenbeck process, $\mathrm{J_c}(r)$. The proof is based on standard results given by Phillips (1987) and, therefore, is omitted.

Theorem 1: *Let y_t be generated by Equation (1) with $\rho = 1$. Let $\hat{\omega}^2(\Delta\hat{x}_t, K)$ be a consistent estimator of the long-run variance for $\hat{x}_t = \{\hat{v}_t, \tilde{v}_t\}$. Then, asymptotically,*

$$\mathrm{CI}^i_{\mathrm{OLS}} \xrightarrow{d} \left\{2\int_0^1 [\bar{V}_i(r)]^2 \mathrm{d}r\right\}^{-1} \tag{5}$$

$$\mathrm{CI}^i_{\mathrm{GLS}} \xrightarrow{d} V_i^{-1} \tag{6}$$

for $i = \mu,\tau$; where $\bar{V}_\mu(s) = \int_0^s W_\mu(u)\mathrm{d}u$, $W_\mu(u) = W(u) - \int W(u)\mathrm{d}u$, $\bar{V}_\tau(s) = \int_0^s W_\tau(u)\mathrm{d}u$, $W_\tau(u) = W_\mu(u) - 12(u-0.5)\int(r-0.5)W(r)dr$. Also, $V_\mu = 2\int \bar{W}(s)^2 ds + (\int W(r)\mathrm{d}r)^2 - 2(\int W(r)\mathrm{d}r)(\int \bar{W}(s)ds)$ with $\bar{W}(s) = \int_0^s W(u)du$, $V_\tau = 2\int[\bar{W}_\tau(s)ds]^2 + (\int W(r)\mathrm{d}r - W_1^*/2a)^2 - 2(\int W(r)dr - W_1^*/2a)(\int \bar{W}_\tau(s)ds)$ with $\bar{W}_\tau(s) = \bar{W}(s) - s^2 W_1^*/2a$, $W_1^* = (1-\bar{c})W(1) + \bar{c}^2\int rW(r)dr$ and $a = 1 - \bar{c} + \bar{c}^2/3$.

3.2 Autoregressive spectral density estimator

Now, we replace the kernel estimator in the numerator with the autoregressive spectral density estimate of ω^2 for a consistent estimator of the long-run variance that no longer depends on $\hat{\rho}$:

$$\hat{\omega}^2_{\mathrm{AR}} = \hat{\sigma}^2_k/(1-\hat{b}(1))^2, \tag{7}$$

where $\hat{b}(1) = \sum_{j=1}^k \hat{b}_j$, $\hat{\sigma}^2_k = T^{-1}\sum_{t=k+1}^T \hat{e}^2_{tk}$, with $\hat{b}_j$ and $\hat{e}_{tk}$ estimated using OLS from the kth order augmented Dickey and Fuller (1979) autoregression:

$$\Delta\hat{x}^i_t = b_0\hat{x}^i_{t-1} + \sum_{j=1}^k b_j \Delta\hat{x}^i_{t-j} + e_{tk}, \tag{8}$$

where k denotes the number of lags to be selected. The consistency of this estimator with OLS-detrended data is demonstrated in Stock (1999), whereas the consistency based on local GLS-detrended data is formalised in Ferrer-Pérez (2016). As we keep the inconsistency under the alternative of the estimator in the denominator, the combination of consistent–inconsistent quotient holds.

Then, combining the use of OLS/GLS detrending procedures with the autoregressive long-run variance estimator in the numerator and the kernel-based estimator in the denominator results in the following statistics:

$$\mathrm{CI}^i_{\mathrm{AR,OLS}} = T^2 \frac{\hat{\omega}^2_{\mathrm{AR,OLS}}}{\hat{\omega}^2(\hat{v}^i_t, T)} \tag{9}$$

$$\mathrm{CI}^i_{\mathrm{AR,GLS}} = T^2 \frac{\hat{\omega}^2_{\mathrm{AR,GLS}}}{\hat{\omega}^2(\tilde{v}^i_t, T)} \tag{10}$$

for $i = \mu,\tau$. The subscript AR indicates that the test uses the autoregressive spectral density estimator in the numerator. In addition, the subscript OLS reflects that both the numerator and the denominator are based on OLS-detrended data, and the subscript

GLS indicates that both the numerator and the denominator are based on GLS-detrended data. Given the arguments above, it can be easily deduced that Theorem 1 also holds for these two modified test statistics.

Critical values are calculated by approximating the Wiener process by partial sums of standard normal random variables with zero-mean and variance equal to unity with 10,000 steps and 100,000 replications. Table 1 summarises these values.

To investigate the asymptotic behaviour we have simulated the local asymptotic distributions of the tests. Figure 1 illustrates the asymptotic power functions for the mean case in Figure 1a and for the trend case in Figure 1b. We have computed Monte Carlo simulations with 50,000 replications.

From Figure 1, we can see that the asymptotic size seems to attain nominal levels for all of the tests and that the GLS-based (and also OLS-based in the trend case) test produces substantial power gains compared to the original test as c tends to $-\infty$; that is, the non-centrality parameter deviates from the unit root null. In the trend case, the power advantage obtained with the CI^{τ}_{GLS} with respect to the CI^{τ}_{OLS} tends to reduce as we move away from the null hypothesis.

4. Finite sample results

In this section, we only discuss the results for the small-sample size and size-adjusted power delivered by the statistics analysed in the previous section. For comparison purposes, we show the results for the original CI test as well.

The data generating process in our simulations is given in Equation (1), with the initial value $v_1 = 0$ for T = (100,500) using 20,000 replications. The error term u_t follows either an AR(1) of the form $u_t = \phi u_{t-1} + e_t$ or an MA(1) of the form $u_t = e_t + \theta e_{t-1}$, where $e_t \sim iidN(0, \sigma^2_e)$ and $\sigma^2_e = 1$. We consider ϕ = (0,−0.8,−0.5,0.5,0.8) and θ = (−0.8,−0.5,0.5,0.8). Size results are computed with ρ = 1 and size-adjusted power with ρ = (0.99,0.95,0.9,0.8,0.5). We set $k_{\min} = 0$ and $k_{\max} = \lfloor 12(T/100)^{0.25} \rfloor$. We report the results for a 5% nominal significance level.

Because the definition of the proposed CI tests require either the selection of the bandwidth parameter or the truncation lag parameter or both, we have considered different selection methods depending on the class of the long-run variance estimator used in the test. Thus, for kernel-based estimators, we have applied the automatic bandwidth selection method suggested by Andrews and Monahan (1992) with AR(1) pre-whitened errors. For autoregressive-based estimators, we have selected the truncation lag parameter k using the the modified Akaike information criterion (MAIC) criterion with GLS

Table 1
Asymptotic critical values†

	99%	97.5%	95%	90%	50%	10%	5%	2.5%	1%
CI^{μ}_{OLS}	7.2	9.5	12.4	17.4	88.9	638.9	1,086.4	1,683.8	2746.5
CI^{μ}_{GLS}	0.6	0.8	1.1	1.5	8.4	88.1	172.9	307.0	584.9
CI^{τ}_{OLS}	110.5	143.7	183.0	249.3	893.0	3,626.7	5,351.0	7,439.2	10742.8
CI^{τ}_{GLS}	8.9	11.6	15.2	21.3	105.0	664.4	1,080.8	1,630.4	2601.8

Notes: †In Cai and Shintani (2006, table 2, pp. 354), the asymptotic critical values for the 10, 5 and 1% for the demeaned CI test are 643, 1100 and 2790, respectively. Differences are trivial.

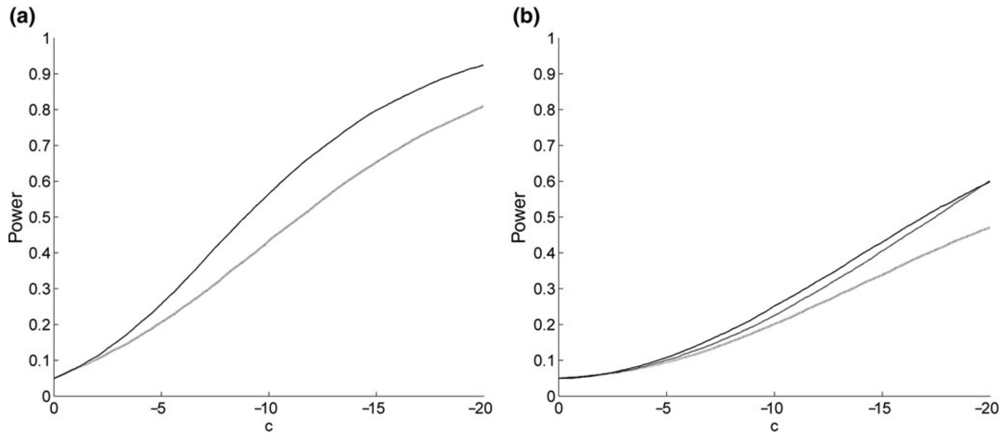

FIGURE 1. Asymptotic size and local power: (a) Mean case and (b) trend case. (—) CI; (········) CI_{OLS}; (—) CI_{GLS}

detrended (demeaned) data and have used that optimal lag parameter for the GLS-based test as in Ng and Perron (2001). The reason why we consider this information criterion instead of the standard criteria such as the Akaike information criterion or the the Bayesian information criterion is because Ng and Perron showed that the use of the MAIC leads to gains in the size-power trade-off of the tests over standard information criteria. However, Perron and Qu (2007) showed that using MAIC with GLS data may lead to power reversal problems of the tests against non-local alternatives. To deal with this issue, they propose a simple solution that consists of using the MAIC criterion with OLS detrended (demeaned) data to select the truncation lag parameter and then using this optimal lag parameter for the GLS-based test. This method is indicated throughout this paper with the subscript PQ.

We can construct six modified test statistics by combining the aforementioned detrending (demeaning) procedures with the two classes of long-run variance estimators. A brief description of these modified CI tests is detailed next (for $i = \mu,\tau$).

CI^i: the original CI test proposed by Cai and Shintani (2006).

CI^i_{OLS}: CI that uses the kernel-based estimator based on OLS-detrended (demeaned) data in both the numerator and the denominator.[7]

CI^i_{GLS}: CI that uses the kernel-based estimator based on GLS-detrended (demeaned) data in both the numerator and the denominator.

$CI^i_{AR,OLS}$: CI that uses the autoregressive spectral density estimator $\hat{\omega}^2_{AR}$ given in Equation (7) in the numerator with OLS-detrended (demeaned) data and the kernel-based estimator in the denominator with OLS detrended (demeaned) data.

$CI^i_{AR,GLS}$: CI that uses the autoregressive spectral density estimator $\hat{\omega}^2_{AR}$ as in Ng and Perron (2001) in the numerator and the kernel-based estimator in the denominator with GLS-detrended (demeaned) data.

$CI^i_{PQ,OLS}$: CI that uses the $\hat{\omega}^2_{AR}$ in the numerator constructed as in Perron and Qu (2007) and the kernel-based estimator in the denominator with OLS-detrended (demeaned) data.

7 Recall that, in the mean case, this version of the test is equivalent to the original CI.

$CI^{i}_{PQ,GLS}$: CI that uses the $\hat{\omega}^2_{AR}$ in the numerator constructed as in Perron and Qu (2007) and the kernel-based estimator in the denominator with GLS-detrended (demeaned) data.

4.1 Empirical size

Tables 2 and 3 show the empirical size. In general, when the errors are iid ($\phi = 0$) the size values are reasonably close to the nominal size, except for those tests that use the GLS procedure, which tends to produce a slightly superior size. In the mean case, for small samples ($T = 100$) the size of the CI^{μ} test is close to the 5%, level, as are those of the $CI^{\mu}_{AR,OLS}$ and $CI^{\mu}_{PQ,OLS}$. In contrast, those based on the GLS procedure are slightly oversized. As the sample size increases ($T = 500$), the tests achieve the 5% nominal size. In the trend case, the picture is quite similar.

Next, we explore the size value of the tests in the presence of serial correlation. When a first-order autoregressive process underlies the errors, Table 2 presents the results for the mean and trend cases. In the mean case, we can see that the original CI^{μ} test is clearly undersized for $\phi > 0$. By applying the local GLS demeaning procedure

TABLE 2
Empirical size, AR(1) errors

Mean case T	ϕ	CI^{μ}	CI^{μ}_{GLS}	$CI^{\mu}_{AR,OLS}$	$CI^{\mu}_{AR,GLS}$	$CI^{\mu}_{PQ,OLS}$	$CI^{\mu}_{PQ,GLS}$
100	0	0.044	0.072	0.038	0.067	0.038	0.067
	0.8	0.010	0.035	0.037	0.070	0.038	0.069
	0.5	0.023	0.051	0.040	0.070	0.039	0.070
	−0.5	0.051	0.080	0.043	0.068	0.041	0.069
	−0.8	0.050	0.081	0.042	0.066	0.039	0.066
500	0	0.052	0.055	0.049	0.055	0.050	0.055
	0.8	0.031	0.041	0.047	0.055	0.047	0.055
	0.5	0.040	0.048	0.048	0.056	0.048	0.056
	−0.5	0.056	0.060	0.050	0.055	0.049	0.055
	−0.8	0.057	0.062	0.051	0.055	0.050	0.055

Trend case T	ϕ	CI^{τ}	CI^{τ}_{OLS}	CI^{τ}_{GLS}	$CI^{\tau}_{AR,OLS}$	$CI^{\tau}_{AR,GLS}$	$CI^{\tau}_{PQ,OLS}$	$CI^{\tau}_{PQ,GLS}$
100	0	0.047	0.040	0.073	0.032	0.063	0.033	0.063
	0.8	0.005	0.002	0.011	0.030	0.060	0.028	0.061
	0.5	0.020	0.013	0.034	0.034	0.064	0.033	0.065
	−0.5	0.044	0.039	0.073	0.045	0.069	0.041	0.070
	−0.8	0.035	0.029	0.058	0.045	0.066	0.041	0.067
500	0	0.052	0.051	0.056	0.046	0.053	0.046	0.053
	0.8	0.025	0.019	0.029	0.042	0.050	0.041	0.050
	0.5	0.037	0.034	0.039	0.048	0.052	0.047	0.052
	−0.5	0.055	0.055	0.060	0.049	0.053	0.048	0.053
	−0.8	0.052	0.052	0.057	0.049	0.053	0.048	0.053

Notes: Results of empirical size at the 5% level are based on asymptotic critical values with data generated under the null hypothesis ($c = 0$). When the autoregressive long-run variance estimator was used in the tests, we followed Ng and Perron (2001) and chose the modified Akaike information criterion to select the truncation lag parameter k. To do this, we set $k_{\min} = 0$ and $k_{\max} = int(12(T/100)^{0.25})$. When the kernel-based long-run variance estimator was used in the tests, we considered the Bartlett kernel function and selected the bandwidth parameter using the automatic selection procedure of Andrews and Monahan (1992) with AR(1) pre-whitened errors. Results are based on 10,000 replications.

TABLE 3
Empirical size, MA(1) errors

Mean case T	θ	CI^{μ}	CI^{μ}_{GLS}	$CI^{\mu}_{AR,OLS}$	$CI^{\mu}_{AR,GLS}$	$CI^{\mu}_{PQ,OLS}$	$CI^{\mu}_{PQ,GLS}$
100	0.8	0.030	0.060	0.047	0.071	0.037	0.069
	0.5	0.032	0.062	0.042	0.071	0.038	0.071
	−0.5	0.088	0.118	0.070	0.087	0.060	0.088
	−0.8	0.334	0.304	0.175	0.114	0.129	0.131
500	0.8	0.045	0.053	0.055	0.058	0.050	0.058
	0.5	0.046	0.052	0.050	0.057	0.048	0.057
	−0.5	0.077	0.080	0.059	0.061	0.055	0.061
	−0.8	0.254	0.208	0.086	0.071	0.071	0.070

Trend case T	θ	CI^{τ}	CI^{τ}_{OLS}	CI^{τ}_{GLS}	$CI^{\tau}_{AR,OLS}$	$CI^{\tau}_{AR,GLS}$	$CI^{\tau}_{PQ,OLS}$	$CI^{\tau}_{PQ,GLS}$
100	0.8	0.027	0.020	0.046	0.041	0.072	0.031	0.070
	0.5	0.030	0.023	0.051	0.038	0.068	0.033	0.068
	−0.5	0.083	0.090	0.125	0.094	0.103	0.082	0.108
	−0.8	0.272	0.478	0.341	0.323	0.175	0.263	0.217
500	0.8	0.043	0.040	0.045	0.060	0.059	0.052	0.058
	0.5	0.044	0.041	0.046	0.050	0.053	0.047	0.053
	−0.5	0.083	0.091	0.085	0.066	0.064	0.061	0.065
	−0.8	0.281	0.397	0.268	0.121	0.078	0.096	0.078

Notes: Results of empirical size at the 5% level are based on asymptotic critical values with data generated under the null hypothesis (c = 0). When the autoregressive long-run variance estimator was used in the tests, we followed Ng and Perron (2001) and chose the modified Akaike information criterion to select the truncation lag parameter k. To do this, we set $k_{\min} = 0$ and $k_{\max} = int(12(T/100)^{0.25})$. When the kernel-based long-run variance estimator was used in the tests, we considered the Bartlett kernel function and selected the bandwidth parameter using the automatic selection procedure of Andrews and Monahan (1992) with AR(1) pre-whitened errors. Results are based on 10,000 replications.

and replacing the kernel-based estimator with the autoregressive estimator, the tests offer improved accuracy and stability, especially for T = 500. In the trend case, we can see that $\hat{\omega}^2_{AR}$-based tests produce improved size compared to the original CI^{τ} test, which suffers from some size distortions for T = 100 and ϕ = (0.8,0.5) as in the case of the kernel-based tests CI^{τ}_{OLS} and CI^{τ}_{GLS}.

In the case of MA(1) errors (Table 3) all of the tests exhibit some over-rejection in the mean case when the parameter θ is negative and close to −1. That is, applying alternative demeaning procedures or replacing the kernel-based estimator in the numerator with the autoregressive estimator does not eliminate the problem completely. However, when θ = −0.8, size distortions are less severe for $\hat{\omega}^2_{AR}$-based ratio tests as they show rejection frequencies that tend to be lower in comparison with those of the CI^{μ} (almost one-third smaller), CI^{μ}_{GLS}. In this case, the $CI^{\mu}_{AR,GLS}$ shows the closest empirical size to the nominal size. The picture is quite similar in the trend case.

4.2 Size-adjusted power

Tables 4–8 report the size-adjusted power. The generating process is as in Equation (1) with T = (100,500) and ρ = (0.99,0.95,0.9,0.8,0.5), which are typical values in the finite sample experiments within the unit root literature.

Table 4
Size-adjusted power, iid errors

Mean case T	ρ	CI^{μ}	CI^{μ}_{GLS}	$CI^{\mu}_{AR,OLS}$	$CI^{\mu}_{AR,GLS}$	$CI^{\mu}_{PQ,OLS}$	$CI^{\mu}_{PQ,GLS}$
100	0.99	0.074	0.077	0.073	0.075	0.073	0.075
	0.95	0.197	0.261	0.195	0.249	0.194	0.250
	0.9	0.418	0.605	0.398	0.554	0.397	0.561
	0.8	0.808	0.959	0.725	0.854	0.721	0.873
	0.5	0.999	1.000	0.889	0.933	0.882	0.949
500	0.99	0.188	0.259	0.185	0.251	0.185	0.252
	0.95	0.890	0.982	0.860	0.962	0.859	0.965
	0.9	0.999	1.000	0.986	0.998	0.985	0.998
	0.8	1.000	1.000	0.999	1.000	0.999	1.000
	0.5	1.000	1.000	1.000	1.000	1.000	1.000

Trend case T	ρ	CI^{τ}	CI^{τ}_{OLS}	CI^{τ}_{GLS}	$CI^{\tau}_{AR,OLS}$	$CI^{\tau}_{AR,GLS}$	$CI^{\tau}_{PQ,OLS}$	$CI^{\tau}_{PQ,GLS}$
100	0.99	0.053	0.056	0.053	0.055	0.053	0.055	0.054
	0.95	0.095	0.105	0.111	0.106	0.109	0.104	0.110
	0.9	0.205	0.237	0.267	0.231	0.256	0.226	0.258
	0.8	0.477	0.629	0.672	0.582	0.601	0.571	0.613
	0.5	0.813	0.996	0.991	0.877	0.832	0.866	0.884
500	0.99	0.092	0.100	0.107	0.099	0.107	0.100	0.108
	0.95	0.572	0.753	0.740	0.717	0.708	0.716	0.713
	0.9	0.829	0.995	0.980	0.967	0.944	0.966	0.952
	0.8	0.962	1.000	1.000	0.998	0.987	0.997	0.991
	0.5	0.998	1.000	1.000	1.000	0.984	1.000	0.991

Notes:Results represent empirical rejection frequencies of 5% level tests based on size-adjusted critical values with data generated under the alternative hypothesis for ρ = (0.99,0.9,0.8,0.5). When the autoregressive long-run variance estimator was used in the tests, we followed Ng and Perron (2001) and chose the modified Akaike information criterion to select the truncation lag parameter k. We set $k_{\min} = 0$ and $k_{\max} = int(12(T/100)^{0.25})$. When the kernel-based long-run variance estimator was used in the tests, we considered the Bartlett kernel function and selected the bandwidth parameter using the automatic selection procedure of Andrews and Monahan (1992) with AR(1) pre-whitened errors. Results are based on 10,000 replications.

Table 4 reports the size-adjusted power for the case of iid errors. As expected, the reported size-adjusted power of all of the tests increases as ρ decreases. In addition, we can see in the mean case that the CI^{μ}_{GLS} test produces, in general, the highest size-adjusted power, even for local-to-unity alternatives, followed by the $CI^{\mu}_{AR,GLS}$ and $CI^{\mu}_{PQ,GLS}$ tests. The size-adjusted power of the CI^{μ} test is slightly lower, followed by the $CI^{\mu}_{AR,OLS}$ and $CI^{\mu}_{PQ,OLS}$. For the trend case and for local alternatives (ρ = 0.99,0.95), all the tests produce a size-adjusted power that seems quite similar in both small and large samples. However, for ρ = 0.90,0.8,0.5, the size-adjusted power reported by the modified tests is usually higher than that obtained with the original test. For example, when ρ = 0.8 and T = 100, the power of the CI^{τ} test is 47.7% and those of the CI^{τ}_{GLS} and CI^{τ}_{OLS} are 67.2 and 62.9%, respectively.

Tables 5 and 6 show, respectively, the size-adjusted power of the tests for T = 100,500 when the errors follow an AR(1) process. Focusing on the case that attracts most interest from researchers, namely, when $\phi > 0$, in small samples, minimal differences can be observed between the local size-adjusted power of the tests in both the mean and trend cases. Modified tests that consider GLS-demeaned data, $\hat{\omega}^2_{AR,GLS}$, and the solution of PQ exhibit slightly superior values even for large positive ϕ values.

TABLE 5
Size-adjusted power, T = 100, AR(1) errors

Mean case ρ	ϕ	CI^{μ}	CI^{μ}_{GLS}	$CI^{\mu}_{AR,OLS}$	$CI^{\mu}_{AR,GLS}$	$CI^{\mu}_{PQ,OLS}$	$CI^{\mu}_{PQ,GLS}$
0.99	0.8	0.069	0.072	0.067	0.071	0.067	0.072
	0.5	0.073	0.074	0.072	0.073	0.071	0.074
	−0.5	0.073	0.076	0.074	0.077	0.074	0.076
	−0.8	0.070	0.076	0.073	0.077	0.074	0.078
0.95	0.8	0.151	0.199	0.152	0.185	0.144	0.187
	0.5	0.178	0.234	0.178	0.226	0.178	0.230
	−0.5	0.183	0.252	0.194	0.255	0.196	0.255
	−0.8	0.169	0.243	0.194	0.260	0.198	0.263
0.9	0.8	0.266	0.386	0.258	0.345	0.248	0.351
	0.5	0.347	0.506	0.337	0.474	0.336	0.480
	−0.5	0.363	0.551	0.391	0.560	0.395	0.566
	−0.8	0.325	0.518	0.388	0.561	0.397	0.571
0.8	0.8	0.468	0.660	0.421	0.560	0.407	0.572
	0.5	0.658	0.860	0.573	0.753	0.574	0.757
	−0.5	0.689	0.897	0.708	0.840	0.711	0.859
	−0.8	0.649	0.877	0.698	0.826	0.703	0.850
0.5	0.8	0.771	0.928	0.593	0.750	0.579	0.754
	0.5	0.974	1.000	0.858	0.925	0.858	0.943
	−0.5	0.976	0.999	0.883	0.935	0.879	0.949
	−0.8	0.985	1.000	0.889	0.924	0.881	0.944

Trend case ρ	ϕ	CI^{τ}	CI^{τ}_{OLS}	CI^{τ}_{GLS}	$CI^{\tau}_{AR,OLS}$	$CI^{\tau}_{AR,GLS}$	$CI^{\tau}_{PQ,OLS}$	$CI^{\tau}_{PQ,GLS}$
0.99	0.8	0.052	0.049	0.052	0.051	0.052	0.052	0.052
	0.5	0.053	0.054	0.055	0.053	0.055	0.054	0.054
	−0.5	0.053	0.054	0.054	0.055	0.054	0.055	0.054
	−0.8	0.054	0.053	0.053	0.055	0.053	0.055	0.053
0.95	0.8	0.085	0.082	0.095	0.075	0.089	0.076	0.090
	0.5	0.090	0.096	0.110	0.094	0.106	0.095	0.105
	−0.5	0.091	0.096	0.109	0.103	0.107	0.104	0.106
	−0.8	0.089	0.093	0.105	0.102	0.107	0.104	0.106
0.9	0.8	0.142	0.153	0.186	0.127	0.160	0.128	0.162
	0.5	0.173	0.195	0.232	0.182	0.215	0.183	0.213
	−0.5	0.177	0.198	0.237	0.224	0.246	0.229	0.248
	−0.8	0.158	0.185	0.220	0.222	0.245	0.226	0.247
0.8	0.8	0.272	0.324	0.372	0.246	0.295	0.246	0.298
	0.5	0.371	0.468	0.530	0.346	0.428	0.352	0.422
	−0.5	0.354	0.488	0.538	0.559	0.572	0.561	0.586
	−0.8	0.302	0.477	0.489	0.555	0.537	0.557	0.558
0.5	0.8	0.522	0.682	0.707	0.366	0.454	0.372	0.446
	0.5	0.722	0.941	0.944	0.765	0.769	0.772	0.779
	−0.5	0.521	0.945	0.862	0.859	0.764	0.846	0.838
	−0.8	0.429	0.974	0.800	0.857	0.658	0.832	0.773

Notes: Results represent empirical rejection frequencies of 5% level tests based on size-adjusted critical values with data generated under the alternative hypothesis for ρ = (0.99,0.9,0.8,0.5). When the autoregressive long-run variance estimator was used in the tests, we followed Ng and Perron (2001) and chose the modified Akaike information criterion to select the truncation lag parameter k. We set $k_{\min} = 0$ and $k_{\max} = int(12(T/100)^{0.25})$. When the kernel-based long-run variance estimator was used in the tests, we considered the Bartlett kernel function and selected the bandwidth parameter using the automatic selection procedure of Andrews and Monahan (1992) with AR(1) pre-whitened errors. Results are based on 10,000 replications.

For example, in the mean case of Table 5, for $\rho = 0.95$ and $\phi = 0.5$, while CI^{μ} gives 0.178, the CI^{μ}_{GLS} and $CI^{\mu}_{AR,GLS}$ deliver 0.234 and 0.226, respectively. Moreover, as we can see in the trend case, our modified tests based on OLS and $\hat{\omega}^2_{AR,OLS}$ usually attain

Table 6
Size-adjusted power, T = 500, AR(1) errors

Mean case ρ	ϕ	CI^{μ}	CI^{μ}_{GLS}	$CI^{\mu}_{AR,OLS}$	$CI^{\mu}_{AR,GLS}$	$CI^{\mu}_{PQ,OLS}$	$CI^{\mu}_{PQ,GLS}$
0.99	0.8	0.175	0.229	0.180	0.234	0.181	0.234
	0.5	0.183	0.242	0.186	0.244	0.186	0.244
	−0.5	0.185	0.248	0.187	0.250	0.187	0.250
	−0.8	0.179	0.242	0.186	0.249	0.185	0.250
0.95	0.8	0.750	0.897	0.778	0.908	0.777	0.910
	0.5	0.832	0.954	0.839	0.948	0.838	0.950
	−0.5	0.843	0.963	0.860	0.963	0.862	0.965
	−0.8	0.799	0.938	0.857	0.960	0.857	0.963
0.9	0.8	0.947	0.993	0.952	0.991	0.951	0.993
	0.5	0.989	1.000	0.979	0.997	0.978	0.997
	−0.5	0.991	1.000	0.986	0.998	0.986	0.998
	−0.8	0.978	0.999	0.985	0.998	0.984	0.998
0.8	0.8	0.996	1.000	0.990	0.999	0.990	0.999
	0.5	1.000	1.000	0.998	1.000	0.998	1.000
	−0.5	1.000	1.000	0.999	1.000	0.999	1.000
	−0.8	1.000	1.000	1.000	1.000	0.999	1.000
0.5	0.8	1.000	1.000	0.998	1.000	0.998	1.000
	0.5	1.000	1.000	1.000	1.000	1.000	1.000
	−0.5	1.000	1.000	1.000	1.000	1.000	1.000
	−0.8	1.000	1.000	1.000	1.000	1.000	1.000

Trend case ρ	ϕ	CI^{τ}	CI^{τ}_{OLS}	CI^{τ}_{GLS}	$CI^{\tau}_{AR,OLS}$	$CI^{\tau}_{AR,GLS}$	$CI^{\tau}_{PQ,OLS}$	$CI^{\tau}_{PQ,GLS}$
0.99	0.8	0.084	0.092	0.101	0.095	0.105	0.094	0.104
	0.5	0.089	0.097	0.110	0.097	0.107	0.098	0.107
	−0.5	0.090	0.097	0.108	0.099	0.108	0.099	0.108
	−0.8	0.087	0.094	0.104	0.099	0.109	0.099	0.109
0.95	0.8	0.428	0.548	0.568	0.588	0.608	0.583	0.609
	0.5	0.505	0.661	0.674	0.672	0.678	0.673	0.680
	−0.5	0.504	0.679	0.681	0.714	0.706	0.715	0.709
	−0.8	0.434	0.615	0.608	0.711	0.696	0.712	0.700
0.9	0.8	0.644	0.883	0.839	0.897	0.865	0.893	0.871
	0.5	0.744	0.971	0.937	0.951	0.925	0.949	0.933
	−0.5	0.710	0.976	0.936	0.967	0.940	0.965	0.949
	−0.8	0.593	0.948	0.868	0.966	0.925	0.963	0.936
0.8	0.8	0.784	0.989	0.964	0.978	0.959	0.975	0.966
	0.5	0.878	1.000	0.997	0.994	0.983	0.993	0.987
	−0.5	0.801	1.000	0.992	0.998	0.983	0.997	0.988
	−0.8	0.653	1.000	0.957	0.998	0.962	0.997	0.976
0.5	0.8	0.912	1.000	0.999	0.995	0.985	0.995	0.989
	0.5	0.983	1.000	1.000	1.000	0.992	1.000	0.997
	−0.5	0.800	1.000	0.997	1.000	0.950	1.000	0.977
	−0.8	0.652	1.000	0.969	1.000	0.839	1.000	0.929

Notes: Results represent empirical rejection frequencies of 5% level tests based on size-adjusted critical values with data generated under the alternative hypothesis for ρ = (0.99,0.9,0.8,0.5). When the autoregressive long-run variance estimator was used in the tests, we followed Ng and Perron (2001) and chose the modified Akaike information criterion to select the truncation lag parameter k. We set $k_{\min} = 0$ and $k_{\max} = int(12(T/100)^{0.25})$. When the kernel-based long-run variance estimator was used in the tests, we considered the Bartlett kernel function and selected the bandwidth parameter using the automatic selection procedure of Andrews and Monahan (1992) with AR(1) pre-whitened errors. Results are based on 10,000 replications.

larger power values compared to the original CI^{τ} as ρ decreases. In large samples, the size-adjusted power attained by each modified test is greater than that produced by the original test.

Tables 7 and 8 report the results when the errors follow an MA(1) process for $T = 100$ and $T = 500$, respectively. We will focus on the case that usually attracts most interest, when $\theta < 0$. In general, we can see that as ρ decreases, the

Table 7
Size-adjusted power, $T = 100$, MA(1) errors

Mean case ρ	θ	CI^{μ}	CI^{μ}_{GLS}	$CI^{\mu}_{AR,OLS}$	$CI^{\mu}_{AR,GLS}$	$CI^{\mu}_{PQ,OLS}$	$CI^{\mu}_{PQ,GLS}$
0.99	0.8	0.072	0.073	0.070	0.073	0.071	0.073
	0.5	0.072	0.074	0.072	0.074	0.073	0.075
	−0.5	0.073	0.076	0.072	0.075	0.074	0.076
	−0.8	0.068	0.079	0.069	0.078	0.072	0.077
0.95	0.8	0.179	0.239	0.164	0.216	0.176	0.216
	0.5	0.184	0.246	0.182	0.232	0.188	0.234
	−0.5	0.186	0.260	0.181	0.250	0.190	0.253
	−0.8	0.145	0.287	0.148	0.277	0.173	0.277
0.9	0.8	0.359	0.525	0.299	0.434	0.329	0.438
	0.5	0.371	0.545	0.343	0.487	0.356	0.492
	−0.5	0.377	0.586	0.355	0.530	0.375	0.542
	−0.8	0.308	0.651	0.324	0.591	0.375	0.593
0.8	0.8	0.685	0.884	0.548	0.721	0.587	0.738
	0.5	0.714	0.908	0.621	0.789	0.639	0.802
	−0.5	0.739	0.935	0.657	0.803	0.674	0.824
	−0.8	0.701	0.966	0.718	0.886	0.762	0.897
0.5	0.8	0.983	1.000	0.828	0.902	0.843	0.924
	0.5	0.994	1.000	0.879	0.928	0.883	0.946
	−0.5	0.993	1.000	0.889	0.935	0.890	0.952
	−0.8	0.995	1.000	0.993	0.982	0.994	0.997

Trend case ρ	θ	CI^{τ}	CI^{τ}_{OLS}	CI^{τ}_{GLS}	$CI^{\tau}_{AR,OLS}$	$CI^{\tau}_{AR,GLS}$	$CI^{\tau}_{PQ,OLS}$	$CI^{\tau}_{PQ,GLS}$
0.99	0.8	0.053	0.053	0.055	0.050	0.053	0.052	0.053
	0.5	0.053	0.053	0.055	0.052	0.054	0.053	0.054
	−0.5	0.054	0.053	0.052	0.054	0.053	0.054	0.053
	−0.8	0.052	0.053	0.053	0.053	0.052	0.053	0.051
0.95	0.8	0.090	0.097	0.111	0.085	0.099	0.091	0.099
	0.5	0.091	0.097	0.109	0.091	0.105	0.093	0.105
	−0.5	0.095	0.097	0.107	0.097	0.108	0.097	0.108
	−0.8	0.086	0.086	0.113	0.090	0.115	0.092	0.114
0.9	0.8	0.177	0.201	0.237	0.165	0.198	0.181	0.201
	0.5	0.182	0.204	0.244	0.182	0.225	0.188	0.225
	−0.5	0.186	0.204	0.245	0.205	0.239	0.212	0.242
	−0.8	0.153	0.178	0.256	0.187	0.271	0.198	0.281
0.8	0.8	0.383	0.494	0.557	0.379	0.429	0.413	0.439
	0.5	0.406	0.520	0.583	0.417	0.481	0.431	0.479
	−0.5	0.373	0.536	0.578	0.506	0.531	0.517	0.558
	−0.8	0.261	0.501	0.538	0.544	0.599	0.567	0.651
0.5	0.8	0.714	0.962	0.953	0.774	0.751	0.791	0.788
	0.5	0.769	0.982	0.976	0.852	0.828	0.856	0.861
	−0.5	0.516	0.976	0.872	0.859	0.733	0.850	0.848
	−0.8	0.320	0.985	0.696	0.992	0.794	0.992	0.944

Notes: Results represent empirical rejection frequencies of 5% level tests based on size-adjusted critical values with data generated under the alternative hypothesis for ρ = (0.99,0.9,0.8,0.5). When the autoregressive long-run variance estimator was used in the tests, we followed Ng and Perron (2001) and chose the modified Akaike information criterion to select the truncation lag parameter k. We set $k_{\min} = 0$ and $k_{\max} = int(12(T/100)^{0.25})$. When the kernel-based long-run variance estimator was used in the tests, we considered the Bartlett kernel function and selected the bandwidth parameter using the automatic selection procedure of Andrews and Monahan (1992) with AR(1) pre-whitened errors. Results are based on 10,000 replications.

Table 8
Size-adjusted power, T = 500, MA(1) errors

Mean case ρ	θ	CI^{μ}	CI^{μ}_{GLS}	$CI^{\mu}_{AR,OLS}$	$CI^{\mu}_{AR,GLS}$	$CI^{\mu}_{PQ,OLS}$	$CI^{\mu}_{PQ,GLS}$
0.99	0.8	0.184	0.244	0.175	0.233	0.182	0.234
	0.5	0.185	0.247	0.185	0.240	0.188	0.240
	−0.5	0.185	0.255	0.183	0.252	0.185	0.253
	−0.8	0.171	0.275	0.170	0.270	0.183	0.269
0.95	0.8	0.839	0.958	0.795	0.927	0.815	0.927
	0.5	0.850	0.964	0.834	0.945	0.840	0.946
	−0.5	0.863	0.973	0.843	0.953	0.849	0.955
	−0.8	0.865	0.987	0.840	0.966	0.865	0.968
0.9	0.8	0.991	1.000	0.970	0.996	0.975	0.997
	0.5	0.994	1.000	0.982	0.998	0.982	0.998
	−0.5	0.996	1.000	0.983	0.998	0.984	0.998
	−0.8	0.998	1.000	0.992	1.000	0.995	1.000
0.8	0.8	1.000	1.000	0.998	1.000	0.998	1.000
	0.5	1.000	1.000	0.999	1.000	0.999	1.000
	−0.5	1.000	1.000	1.000	1.000	1.000	1.000
	−0.8	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.8	1.000	1.000	1.000	1.000	1.000	1.000
	0.5	1.000	1.000	1.000	1.000	1.000	1.000
	−0.5	1.000	1.000	1.000	1.000	1.000	1.000
	−0.8	1.000	1.000	1.000	1.000	1.000	1.000

Trend case ρ	θ	CI^{τ}	CI^{τ}_{OLS}	CI^{τ}_{GLS}	$CI^{\tau}_{AR,OLS}$	$CI^{\tau}_{AR,GLS}$	$CI^{\tau}_{PQ,OLS}$	$CI^{\tau}_{PQ,GLS}$
0.99	0.8	0.089	0.097	0.109	0.091	0.104	0.095	0.104
	0.5	0.091	0.097	0.108	0.097	0.108	0.098	0.108
	−0.5	0.089	0.097	0.109	0.096	0.107	0.098	0.107
	−0.8	0.090	0.091	0.104	0.092	0.104	0.095	0.104
0.95	0.8	0.510	0.672	0.679	0.601	0.631	0.629	0.631
	0.5	0.522	0.686	0.692	0.665	0.672	0.673	0.673
	−0.5	0.509	0.706	0.699	0.678	0.675	0.689	0.679
	−0.8	0.412	0.715	0.656	0.658	0.645	0.696	0.654
0.9	0.8	0.744	0.975	0.942	0.927	0.902	0.935	0.906
	0.5	0.762	0.981	0.952	0.956	0.930	0.956	0.937
	−0.5	0.715	0.987	0.949	0.959	0.920	0.958	0.928
	−0.8	0.519	0.992	0.864	0.972	0.870	0.976	0.886
0.8	0.8	0.870	1.000	0.997	0.994	0.980	0.994	0.983
	0.5	0.894	1.000	0.999	0.996	0.986	0.996	0.990
	−0.5	0.803	1.000	0.993	0.998	0.970	0.997	0.978
	−0.8	0.551	1.000	0.910	1.000	0.885	1.000	0.954
0.5	0.8	0.963	1.000	1.000	1.000	0.988	1.000	0.992
	0.5	0.985	1.000	1.000	1.000	0.989	1.000	0.994
	−0.5	0.805	1.000	0.997	1.000	0.901	1.000	0.961
	−0.8	0.538	1.000	0.908	1.000	0.702	1.000	1.000

Notes: Results represent empirical rejection frequencies of 5% level tests based on size-adjusted critical values with data generated under the alternative hypothesis for ρ = (0.99,0.9,0.8,0.5). When the autoregressive long-run variance estimator was used in the tests, we followed Ng and Perron (2001) and chose the modified Akaike information criterion to select the truncation lag parameter k. We set $k_{\min} = 0$ and $k_{\max} = int(12(T/100)^{0.25})$. When the kernel-based long-run variance estimator was used in the tests, we considered the Bartlett kernel function and selected the bandwidth parameter using the automatic selection procedure of Andrews and Monahan (1992) with AR(1) pre-whitened errors. Results are based on 10,000 replications.

size-adjusted power of the tests increases more significantly for most modified tests. Moreover, the picture is quite similar to that described in the AR(1) case. When $T = 100$, for local values of ρ, say $\rho = 0.99$, the results of our modified tests are quite similar to those of the original CI test but we start to see improvements in those tests based on GLS-demeaned (detrended) data and also in those that use $\hat{\omega}^2_{\text{AR,GLS}}$ and the solution of PQ. For non-local alternatives, our modified tests tend to exhibit large size-adjusted power values for $\theta < 0$ and most cases of $\theta > 0$, that even double those obtained by the original CI test. For example, in the mean case of Table 7 for $\rho = 0.99$, when $\theta = -0.5$, CI^{μ} produces 0.068, the tests $\text{CI}^{\mu}_{\text{GLS}}, \text{CI}^{\mu}_{\text{AR,OLS}}, \text{CI}^{\mu}_{\text{PQ,OLS}}$ and $\text{CI}^{\mu}_{\text{PQ,GLS}}$ give 0.076, 0.072, 0.072 and 0.077, respectively. For non-local alternatives, say $\rho = 0.9$, when $\theta = -0.8$, the test $\text{CI}^{\mu}(0.308)$ is slightly surpassed by the $\text{CI}^{\mu}_{\text{AR,OLS}}(0.324)$ and $\text{CI}^{\mu}_{\text{PQ,OLS}}(0.375)$ but doubled by the tests $\text{CI}^{\mu}_{\text{GLS}}(0.651), \text{CI}^{\mu}_{\text{AR,GLS}}(0.591)$ and $\text{CI}^{\mu}_{\text{PQ,GLS}}(0.593)$.

In large samples ($T = 500$), the discussion of the results is quite similar but former differences among tests have increased in favour of our modified tests, especially for those that use OLS, $\hat{\omega}^2_{\text{AR,OLS}}$ and the solution of PQ in the trend case, for which their size-adjusted power even doubles that achieved with the original CI^{τ} test. For example, in the trend case of Table 8 when $\rho = 0.9$ and $\theta = -0.8$, the tests $\text{CI}^{\tau}_{\text{OLS}}(0.992)$, $\text{CI}^{\tau}_{\text{GLS}}(0.864), \text{CI}^{\tau}_{\text{AR,OLS}}(0.972), \text{CI}^{\tau}_{\text{AR,GLS}}(0.870), \text{CI}^{\tau}_{\text{PQ,OLS}}(0.976)$ and $\text{CI}^{\tau}_{\text{PQ,GLS}}(0.886)$ exhibit an improved performance compared to the original test $\text{CI}^{\tau}(0.519)$.

5. Conclusions

Cai and Shintani (2006) proposed four unit root tests as a quotient of two long-run variance kernel-based estimators whose rates of convergence under the null and the alternative hypotheses differ. In particular, the combination of a consistent estimator under the null and an inconsistent estimator under the alternative produced the best properties. This test is the CI test.

In this paper, we have modified the CI test with the objective of improving the properties of this test. To do this, we have considered a twofold strategy. First, we have applied the OLS and the GLS of Elliott *et al.* (1996) procedures to remove the deterministic component of the series instead of the particular method considered in the seminal paper. Second, we have replaced the kernel-based estimator in the numerator of the ratio test with the spectral density estimator of the long-run variance, $\hat{\omega}^2_{\text{AR}}$, based on OLS and GLS-detrended (demeaned) data, while keeping the inconsistent kernel-based estimator in the denominator of the ratio test.

We have derived their respective asymptotic behaviour, tabulated new critical values and showed that these new tests clearly produce power gains. Through Monte Carlo simulations, we have examined the finite sample performance of these tests in terms of size and size-adjusted power. To do so, we have constructed a battery of modified test statistics whose results have been compared to those of the original CI test so as to assess the quality of our proposals.

From our results, we can conclude that our proposals mostly lead to tests with an improved size-power trade-off compared to the original CI test in the presence of autocorrelated errors, especially those based on GLS-detrended (demeaned) data followed by those that use the autoregressive spectral density estimator in the numerator.

Acknowledgements

We thank the editor and two anonymous referees for helpful comments. We also wish to thank participants of the IV_t Time Series Econometrics Workshop (U. Zaragoza) for useful comments on an earlier draft. Financial support is acknowledged from the MEC-MICINN under grants ECO2015-65582-DT and ECO2016-74940-P, and from the consolidated research group of "Selección y simulación de modelos econométricos" financed by the Department of Science, Technology and Universities of the Aragonesse Government.

Appendix I

Proofs

We will use the following general notation. Convergence in probability is denoted by $\xrightarrow{p}$ whereas weak convergence in distribution is represented by $\xrightarrow{d}$ as the sample size grows to ∞. $W(r)$ is the standard Wiener process defined on $\mathcal{C}[0,1]$. The integration sign $\int$ denotes integration between 0 and 1. Differencing and lag operators are denoted by Δ and L, respectively; that is, for a stochastic process x_t, we define $\Delta x_t = x_t - x_{t-1}$ and $Lx_t = x_{t-1}$.

Proof. We have that the series y_t is generated by (1) with $d_t = \delta' z_t, z_t = (1,t)'$ and $\rho = 1$. Then, consider the OLS estimates $\hat{\delta}$ obtained from regressing y_t on z_t. We then define the OLS residual as $\hat{v}_t = y_t - \hat{\delta}' z_t$ and its limiting distribution is given by $T^{-1/2}\hat{v}_t \xrightarrow{d} \omega W_\tau(u)$, where $W_\tau(u) = W_\mu(u) - 12(u-1/2)\int(r-1/2)W(r)\mathrm{d}r$. Taking $S_t = \sum_{j=1}^t \hat{v}_j$ and $S_T = 0$, then applying Lemma 1 in Cai and Shintani (2006), we have that $T^{-3/2}\sum_{i=1}^t \hat{v}_i \xrightarrow{d} \omega \int_0^s W_\tau(u)\mathrm{d}u \equiv \omega \bar{V}_\tau(s), T^{-3}S_t^2 \xrightarrow{d} \omega^2 \bar{V}_\tau(s)^2, T^{-4}\sum^T S_t^2 \xrightarrow{d} \omega^2 \int \bar{V}_\tau(s)^2 ds$, and, hence,

$$T^{-2}\hat{\omega}^2(\hat{v}_t, T) = 2T^{-4}\sum^T S_t^2 \xrightarrow{d} 2\omega^2 \int \bar{V}_\tau(s)^2 ds.$$

The proof for the demeaned model is straightforward.

Consider the local-GLS procedure of Elliott *et al.* (1996). Define the GLS residuals as $\tilde{v}_t = y_t - \tilde{\delta}' z_t$. Apply the aforementioned Lemma 1 to $x_t = \tilde{v}_t$. In this case, $S_T = 0$ no longer holds. Thus, in the mean case, the long-run variance estimator is

$$\begin{aligned} T^{-2}\tilde{\omega}^2(\tilde{v}_t, T) &= 2T^{-4}\sum^T S_t^2 + T^{-3}S_T^2 - 2T^{-4}S_T\sum^T S_t \\ &\xrightarrow{d} 2\omega^2 \int \bar{W}(s)^2 ds + \left(\omega \int W(r)\mathrm{d}r\right)^2 \\ &- 2\omega^2\left(\int W(r)\mathrm{d}r\right)\left(\int \bar{W}(s)ds\right). \end{aligned}$$

The null limiting distribution of $\mathrm{CI}_{\mathrm{GLS}}^{\mu}$ follows straightforwardly.

The detrended case follows similar derivations. Following Elliott *et al.* (1996), we obtain that $(\tilde{\delta}_0 - \delta_0)$ is $O(1)$ and $(\tilde{\delta}_1 - \delta_1)$ is $O(T^{-1/2})$ and, hence, the null asymptotic distribution is given by $T^{-1/2}\tilde{v}_t \xrightarrow{d} \omega(W(s) - a^{-1}W_1^*)$, where $a = 1 - \bar{c} + \bar{c}^2/3$ and $W_1^* = (1-\bar{c})W(1) + \bar{c}^2 \int rW(r)\mathrm{d}r$. So, applying the aforementioned Lemma 1 yields

$$T^{-3/2}S_t \xrightarrow{d} \omega \int_0^s W(u)\mathrm{d}u - \omega(s^2/2a)W_1^* = \omega\tilde{W}(s) - \omega(s^2/2a)W_1^* \equiv \omega\bar{W}_\tau(s).$$

The asymptotic behaviour of the long-run variance estimator is

$$\begin{aligned} T^{-2}\tilde{\omega}^2(\tilde{v}_t, T) &= T^{-2}\left[2T^{-2}\sum^{T} S_t^2 + T^{-1}S_T^2 - 2T^{-2}S_T\sum^{T} S_t\right] \\ &\xrightarrow{d} \omega^2\left\{2\int \bar{W}_\tau(s)^2 ds + \left[\int W(r)\mathrm{d}r - W_1^*/(2a)\right]^2\right. \\ &\left. -2\left[\int W(r)\mathrm{d}r - W_1^*/(2a)\right]\left[\int \bar{W}_\tau(s)ds\right]\right\} \equiv \omega^2 V_\tau. \end{aligned}$$

Then, the null asymptotic distribution of $\mathrm{CI}^{\tau}_{\mathrm{GLS}}$ follows directly. ■

Final version accepted 16 April 2018.

REFERENCES

Andrews, D. and J. Monahan (1992) "An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator", *Econometrica* Vol. 60, No. 4, pp. 953–966.

Berk, K. N. (1974) "Consistent Autoregressive Spectral Estimates", *Annals of Statistics*, Vol. 2, No. 3, pp. 489–502.

Cai, Y. and M. Shintani (2006) "On the Alternative Long-run Variance Ratio Test for a Unit Root", *Econometric Theory*, Vol. 22, No. 3, pp. 347–372.

Dickey, D. and W. Fuller (1979) "Distribution of the Estimators for Autorregressive Time Series with a Unit Root", *Journal of the American Statistical Association*, Vol. 74, No. 366, pp. 427–431.

Elliott, G. (1999) "Efficient Tests for a Unit Root When the Initial Observation is Drawn from Its Unconditional Distribution", *International Economic Review*, Vol. 40, No. 3, pp. 767–783.

——— and U. Müller (2006) "Minimizing the Impact of the Initial Condition on Testing for Unit Roots", *Journal of Econometrics*, Vol. 135, No. 1–2, pp. 285–310.

———, T. Rothenberg and J. Stock (1996) "Efficient Tests for an Autoregressive Unit Root", *Econometrica*, Vol. 64, No. 4, pp. 813–836.

Ferrer-Pérez, H. (2016) "El contraste de raíz unitaria MSB: la influencia de la observación inicial" (PhD thesis). University of Zaragoza, Zaragora.

Haldrup, N. and M. Jansson (2006) "Improving Size and Power in Unit Root Testing", in T. Mills and K. Patterson, eds, *Palgrave Handbook of Econometrics*, Vol. 1, New York: Palgrave Macmillan.

Harvey, D. and S. Leybourne (2006) "Power of a Unit-Root Test and the Initial Condition", *Journal of Time Series Analysis*, Vol. 27, No. 5, pp. 739–752.

———, S. Leybourne and A. Taylor (2009) "Unit Root Testing In Practice: Dealing With Uncertainty Over The Trend And Initial Condition", *Econometric Theory*, Vol. 25, No. 3, pp. 587–636.

Kiefer, N. and T. Vogelsang (2002) "Heteroskedasticity-autocorrelation Robust Standard Errors using the Bartlett Kernel without Truncation", *Econometrica*, Vol. 70, No. 5, pp. 2093–2095.

Newey, W. and K. West (1987) "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix", *Econometrica*, Vol. 55, No. 3, pp. 703–708.

Ng, S. and P. Perron (2001) "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power", *Econometrica*, Vol. 69, No. 6, pp. 1519–1554.

Patterson, K. (2011) *Unit Root Tests in Time Series*, Vol. 1, New York: Palgrave Macmillian.

Perron, P. and Z. Qu (2007) "A Simple Modification to Improve the Finite Sample Properties of Ng and Perron's Unit Root Tests", *Economics Letters*, Vol. 94, No. 1, pp. 12–19.

Phillips, P. (1987) "Towards a Unified Asymptotic Theory for Autoregression", *Biometrika*, Vol. 74, No. 3, pp. 535–547.

Stock, J. (1999) "A Class of Tests for Integration and Cointegration", in R. F. Engle and H. White, eds, *Cointegration, Causality, and Forecasting: Festschrift in Honour of Clive W. J. Granger*, pp. 135–167, Oxford: Oxford University Press.