

A comparison of two modified stationarity tests. A Monte Carlo study

H. Ferrer-Pérez^a, M.I. Ayuda^{b,*} and A. Aznar^b

^a*Center for Agro-Food Economy and Development (CREDA-UPC-IRTA), Edifici ESAB Parc Mediterrani de la Tecnologia, 08860, Barcelona, Spain.*

^b*Department of Economic Analysis. University of Zaragoza. Gran Vía 2. 50005-Zaragoza. Spain.*

*Corresponding author: M.I. Ayuda. Tel.: 0034976762410, Fax: 0034976761996.

Abstract: To specify an econometric model with time series data, it is important to determine the order of integration of the variables in the model. In this paper, using a complete set of Monte Carlo experiments, we compare the behaviour of two stationarity tests, the Xiao test and the KPSS (Kwiatkowski, Phillips, Schmidt and Shin) test, using an alternative estimator of the long-run variance to those used in the original version of the tests, to recommend which one to use in practice. First, we compare the small sample properties of the original S_n test with those of its modified version. We conclude that this modified version has a better size versus power trade-off than the original test. So, second, we compare the finite sample properties of the modified S_n and the modified KPSS. Since the modified KPSS exhibits higher power and size, we conduct a second experiment determining the critical value of each test, in such a way that the power of both tests coincides at 0.5, and then we examine their size for some local-to-unity values. The results show that, in most cases, the performance of the modified KPSS test dominates that of the modified S_n test.

Key Words: Stationarity test, KPSS test, S_n test, Monte Carlo experiment, Long-run variance.

JEL classifications: C12, C15, C22

1. Introduction

Testing stationarity in time series econometrics has received a great deal of attention in the last three decades and the debate as to whether economic time series are best characterized as stationary or integrated processes is still open.

In order to discriminate between stationary and integrated processes, two broad approaches can be followed: unit root tests and stationarity tests. In this paper, we focus on stationarity tests.

We compare two of the several stationarity tests (see for example [6, 7, 9, 10, 13] proposed in the literature, namely the KPSS test developed in [6] and the S_n test proposed in [13]. Their corresponding asymptotic distributions can be found in the aforementioned works, to which the interested reader is referred.

A widely accepted result in the literature is that both stationarity tests are oversized in the presence of near integrated processes and it is also well known that overrejection of the null hypothesis depends on the estimator of the long-run variance used. Recently, [3] compare the performance of two stationarity tests using different long-run variance estimators and conclude that the method proposed in [12] outperforms other considered choices.

Therefore, the objective of this paper is twofold. Using a complete set of Monte Carlo experiments, we first evaluate whether the modified S_n test, S_nM henceforth, with the boundary condition rule proposed in [12], when estimating the long-run variance parameter using an automatic bandwidth selection procedure, improves the behaviour of the original S_n test. Then, as this is true, in the second part of the paper, we compare two modified versions of the KPSS and the S_n tests, the KPSSM and the S_nM , respectively, and conclude which of

the two has the best trade-off between size and power, because researchers should aim to use a test with as good a size versus power trade-off as possible.

The paper is organized as follows. Section 2 introduces the KPSS and the S_n tests, both with the new modification proposed by [12]. In Section 3, we firstly present an evaluation of the finite sample size and power properties of the original S_n test and the S_nM test following the boundary condition in [12] and, secondly, we compare the behaviour of the version of the S_n test that has better finite sample properties, with the KPSSM. Finally, some concluding comments are provided in Section 4.

2. The modified KPSS and S_n tests

As in [2, 4, 8, 13], we use the following data generating process (DGP, hereafter) that is standard in the literature:

$$\begin{aligned} y_t &= \mu + \delta t + u_t & t &= 1, \dots, T \\ u_t &= \rho u_{t-1} + v_t \end{aligned} \tag{1}$$

where $\mu + \delta t$ is the deterministic part of the model. We consider three cases: Model 1, without deterministic terms ($\mu = \delta = 0$), Model 2, with a constant ($\mu \neq 0, \delta = 0$) and Model 3, with a constant and a linear trend. We assume that the disturbances v_t follow a stationary process whose coefficients satisfy the summability condition given in the assumptions of [13].

Our interest in this paper, therefore, focuses on testing the null hypothesis that y_t is stationary, $H_0 : |\rho| < 1$, against the unit root alternative $H_1 : \rho = 1$.

The modified version of the KPSS test, KPSSM, is the test proposed in [6] but estimating the long-run variance parameter as in [3] with the boundary condition rule¹ proposed in [12]:

$$KPSSM = \frac{T^{-2} \sum_{t=1}^T S_t^2}{\hat{\lambda}_u^2} \quad (2)$$

where $S_t = \sum_{j=1}^t \hat{u}_j$ is the partial sum process of the OLS residuals, defined as

$\hat{u}_t = y_t - \hat{\mu} - \hat{\delta}t$ where $\hat{\mu}$ and $\hat{\delta}$ are the OLS estimators of the coefficients in the deterministic part, and $\hat{\lambda}_u$ denotes the square root of the estimation of the long-run variance.

The modified version of Xiao's test, S_nM , is the test proposed in [13] with the same estimation of the long-run variance as in Eq. (2):

$$S_nM = \max_{k=1, \dots, T} \frac{k}{\hat{\lambda}_u T^{1/2}} \left| \frac{1}{k} S_k - \frac{1}{T} \sum_{t=1}^T \hat{u}_t \right| \quad (3)$$

The nonparametric estimation of the long-run variance entails the stages below.

First, estimate the following AR(p) model:

$$\hat{u}_t = \phi_1 \hat{u}_{t-1} + \phi_2 \hat{u}_{t-2} + \dots + \phi_p \hat{u}_{t-p} + \varepsilon_t \quad (4)$$

and obtain the prewhitened OLS residuals, denoted as $\hat{\varepsilon}_t$, where a statistical criterion is used to determine the value of p .

Then, the estimator of the long-run variance is defined as:

¹ In our study, we have followed [3] when applying the boundary rule proposed in [12], and we have applied their method to the three models, without a constant term, with a constant term and with a linear trend. [12] point out that they do not provide bias reduction methods for the case where a linear trend is fitted and that this issue needs further research, but it is beyond the objective of this paper.

$$\hat{\lambda}_\varepsilon^2 = T^{-1} \sum_{t=1}^T \hat{\varepsilon}_t^2 + 2T^{-1} \sum_{s=1}^{\ell} \omega(s, \ell) \sum_{t=1}^{\ell} \hat{\varepsilon}_t \hat{\varepsilon}_{t-s} \quad (5)$$

where $\omega(s, \ell)$ denotes the kernel function and ℓ is the bandwidth parameter chosen automatically following [1]. In a second stage, the estimated long-run variance based on $\hat{\varepsilon}_t$ is recoloured²:

$$\hat{\lambda}_u^2 = \frac{\hat{\lambda}_\varepsilon^2}{\hat{\phi}(1)^2} \quad (6)$$

where, $\hat{\phi}(1) = 1 - \hat{\phi}_1 - \dots - \hat{\phi}_p$. Finally, [12] estimate the long-run variance using the following boundary rule to ensure that both the KPSSM and S_nM tests are consistent:

$$\hat{\lambda}_u^2 = \min \left\{ T \hat{\lambda}_\varepsilon^2, \frac{\hat{\lambda}_\varepsilon^2}{\hat{\phi}(1)^2} \right\} \quad (7)$$

3. Size and power of the modified tests in finite samples

The Monte Carlo study that we have carried out in this paper has two objectives. First, to see if the proposed S_nM test based on the estimator of the long-run variance developed in [12] improves the behaviour of the S_n test of [13]. Second, to compare the version of the S_n test that has a better size versus power trade-off, with the KPSSM test, and conclude which of them has the best small sample properties.

The DGP used in the experiments is as in (1):

² When the original S_n test is computed, no colouring method is implemented to deal with the autocorrelation when estimating the long-run variance, as in the original paper. We have designed this strategy to compare the original test in Xiao (2001) -S_n test- with the modified version (S_nM).

$$\begin{aligned} y_t &= \mu + \delta t + u_t & t = 1, \dots, T \\ u_t &= \rho u_{t-1} + v_t \end{aligned}$$

where we consider three different processes for v_t : i.i.d process, a first-order autoregressive process, $v_t = \phi v_{t-1} + \varepsilon_t$ and a first-order moving average process, $v_t = \varepsilon_t - \theta \varepsilon_{t-1}$ with $\varepsilon_t \sim i.i.dN(0,1)$ and several values of the parameter $\phi = (0.2, 0.5, 0.8)$ and $\theta = (0.2, 0.5, 0.8)$.

As indicated in the previous section, we consider three leading cases of the deterministic components, Model 1, Model 2 and Model 3. In both tests, we compute the demeaned and detrended OLS method. The values considered for ρ are 0.1, 0.85, 0.90, 0.95, 0.97, 0.99 and 1. Notice that values close to unity are convenient to investigate because many macroeconomic time series contain a large autoregressive coefficient, and they are also commonly used in the unit root Monte Carlo simulations in the literature. Three different sample sizes are used, $T=50, 100$ and 500 . This study is based on 20,000 replications. The nominal size is set at the 5% level of significance. For the kernel function, following [6] and [13], we consider the Bartlett kernel³. The BIC criterion⁴ of [11] is employed to select the lag length in Eq. (2).

In a first experiment, we compare the S_nM test with three versions of the S_n test - $S_n(4)$, $S_n(6)$ and $S_n(12)$ - where the long-run variance has been estimated using three bandwidth parameters⁵, $l4 = \text{int}[4(T/100)^{1/4}]$, $l6 = \text{int}[6(T/100)^{1/4}]$ and $l12 = \text{int}[12(T/100)^{1/4}]$, respectively, where “int[.]” means the integer value.

³ We have also conducted this study using the Quadratic Spectral kernel, but the results, which did not change, are available from the authors upon request.

⁴ The AIC criterion was also used in all the Monte Carlo experiments presented in this section. However, the results were similar to the ones based on the BIC criterion and, therefore, are omitted. These results are available from the authors upon request.

⁵ The bandwidth parameter corresponds to the number of lags used to calculate the long-run variance and, as [13] concludes from his simulations, in cases of large ρ values; the over-

In Table 1, we provide the empirical size ($\rho = 0.1, 0.85, 0.9, 0.95, 0.97, 0.99$) and power ($\rho = 1$) of the $S_n(4)$, the $S_n(6)$, the $S_n(12)$, and the S_nM tests, for the three models (Model 1, Model 2 and Model 3) and three sample sizes ($T = 50, 100$ and 500), when v_t is an i.i.d sequence. Because the conclusions of the results of the autoregressive and moving average disturbances cases are similar to those for the i.i.d. errors, and for reasons of space, we do not present the whole set of results, which are available from the authors upon request.

Table 1: Size and power of the S_n and S_nM tests, iid errors

T	ρ	Model 1 ($\mu = 0, \delta = 0$)				Model 2 ($\mu = 1, \delta = 0$)				Model 3 ($\mu = 1, \delta = 1$)			
		$S_n(4)$	$S_n(6)$	$S_n(12)$	S_nM	$S_n(4)$	$S_n(6)$	$S_n(12)$	S_nM	$S_n(4)$	$S_n(6)$	$S_n(12)$	S_nM
50	0.1	0.013	0.004	0.000	0.020	0.016	0.009	0.002	0.025	0.007	0.009	0.086	0.035
	0.85	0.098	0.006	0.000	0.002	0.183	0.021	0.000	0.006	0.032	0.001	0.004	0.008
	0.9	0.124	0.007	0.000	0.004	0.252	0.032	0.000	0.005	0.045	0.000	0.002	0.006
	0.95	0.131	0.007	0.000	0.039	0.347	0.051	0.000	0.042	0.056	0.000	0.002	0.008
	0.97	0.127	0.009	0.000	0.077	0.407	0.076	0.000	0.079	0.063	0.000	0.002	0.010
	0.99	0.107	0.004	0.000	0.160	0.504	0.117	0.000	0.163	0.070	0.000	0.002	0.012
	1	0.091	0.004	0.000	0.230	0.566	0.138	0.000	0.232	0.067	0.000	0.001	0.010
100	0.1	0.022	0.014	0.003	0.024	0.028	0.021	0.006	0.026	0.021	0.016	0.018	0.022
	0.85	0.193	0.078	0.001	0.001	0.250	0.128	0.006	0.002	0.193	0.042	0.002	0.001
	0.9	0.264	0.122	0.001	0.002	0.356	0.208	0.010	0.002	0.286	0.067	0.001	0.001
	0.95	0.324	0.177	0.003	0.048	0.508	0.344	0.016	0.049	0.423	0.124	0.001	0.001
	0.97	0.319	0.182	0.004	0.128	0.585	0.415	0.025	0.129	0.480	0.156	0.001	0.008
	0.99	0.285	0.169	0.003	0.283	0.685	0.528	0.041	0.285	0.545	0.202	0.000	0.026
	1	0.265	0.150	0.001	0.440	0.769	0.637	0.047	0.439	0.559	0.212	0.000	0.032
500	0.1	0.040	0.034	0.029	0.036	0.042	0.038	0.034	0.037	0.036	0.035	0.026	0.033
	0.85	0.336	0.190	0.067	0.014	0.347	0.207	0.082	0.018	0.435	0.240	0.068	0.005
	0.9	0.474	0.311	0.103	0.007	0.503	0.340	0.125	0.011	0.627	0.390	0.108	0.002
	0.95	0.702	0.521	0.231	0.004	0.746	0.582	0.276	0.004	0.877	0.681	0.273	0.001
	0.97	0.823	0.658	0.347	0.085	0.879	0.735	0.424	0.084	0.948	0.831	0.424	0.026
	0.99	0.861	0.743	0.463	0.539	0.968	0.906	0.672	0.540	0.991	0.949	0.700	0.366
	1	0.817	0.694	0.399	0.861	0.992	0.976	0.875	0.863	0.991	0.964	0.765	0.560

rejection is severe for both tests when the bandwidth is small. This is because, according to the asymptotic theory, the validity of the tests requires the bandwidth to increase with T in these circumstances. In our simulation study, we have used three bandwidths, in order to see the behaviour of the tests as the bandwidth changes.

In Table 1, we can observe that the values of size and power of the tests decrease as the bandwidth increases, as in [6] and [13]. This trade-off between the choice of bandwidth and the rejecting frequency of the tests statistic is explained as follows: when the sample size is small, if the bandwidth selected is too large, the long run variance is overestimated and the value of the test statistic becomes too small and, hence, the test will have the smallest rejection rates. This inverse relation was previously pointed out by [5].

As we can see from Table 1, especially for Model 1 and for all sample sizes, the S_nM test exhibits a size closer to the nominal size than that of the S_n test, for values of ρ closer to 1. In general, and in terms of power ($\rho = 1$), the S_nM test has a higher power than the original test. For Model 2, $S_n(4)$ and $S_n(6)$ present severe size distortions and, the S_nM presents a size closer to the nominal size, for a higher value of ρ , than the $S_n(12)$. Besides, the S_nM has a higher power. For Model 3, $S_n(4)$ and $S_n(6)$ also have high size distortions and the behaviour of the $S_n(12)$ and the S_nM tests are similar in small samples ($T = 50, 100$). But the result is not so clear when the sample size is large ($T = 500$) because the S_nM has a size close to the nominal for a higher value of ρ but, on the contrary, has less power. So, in general, we observe that the size and power trade-off is better for the S_nM test.

Having accepted that the S_nM test has better small sample properties than the original S_n test, in a second Monte Carlo experiment, we compare the behaviour of this modified test to the KPSS test with the same modification in the estimation of the long-run variance in order to decide which test should be used in applied research.

In the following three tables, we provide the empirical size and power of the KPSSM and S_nM tests for the three models and for the three processes of the disturbance term previously presented. The BIC is employed to select the lag length.

Table 2 shows the results when v_t is a white noise. Tables 3 and 4 show the results for an autoregressive, and a moving average process respectively. For reasons of space, in Tables 3 and 4, we present the results for an intermediate value of the autoregressive and moving average parameter ($\phi, \theta = 0.5$).

Table 2: Size and power, iid errors

T	ρ	Model 1 ($\mu = 0, \delta = 0$)		Model 2 ($\mu = 1, \delta = 0$)		Model 3 ($\mu = 1, \delta = 1$)	
		KPSSM	S_nM	KPSSM	S_nM	KPSSM	S_nM
50	0.1	0.040	0.020	0.045	0.025	0.065	0.035
	0.85	0.021	0.002	0.021	0.006	0.027	0.008
	0.9	0.087	0.004	0.041	0.005	0.030	0.006
	0.95	0.277	0.039	0.117	0.042	0.048	0.008
	0.97	0.382	0.077	0.185	0.079	0.061	0.010
	0.99	0.511	0.160	0.310	0.163	0.072	0.012
	1	0.570	0.230	0.390	0.232	0.077	0.010
100	0.1	0.043	0.024	0.042	0.026	0.055	0.022
	0.85	0.017	0.001	0.013	0.002	0.015	0.001
	0.9	0.031	0.002	0.016	0.002	0.011	0.001
	0.95	0.222	0.048	0.116	0.049	0.055	0.001
	0.97	0.378	0.128	0.228	0.129	0.103	0.008
	0.99	0.599	0.283	0.401	0.285	0.166	0.026
	1	0.702	0.440	0.555	0.439	0.179	0.032
500	0.1	0.051	0.036	0.048	0.037	0.045	0.033
	0.85	0.043	0.014	0.039	0.018	0.035	0.005
	0.9	0.037	0.007	0.029	0.011	0.030	0.002
	0.95	0.040	0.004	0.029	0.004	0.025	0.001
	0.97	0.174	0.085	0.139	0.084	0.121	0.026
	0.99	0.652	0.539	0.602	0.540	0.547	0.366
	1	0.917	0.861	0.881	0.863	0.716	0.560

Table 3: Size and power, AR(1) errors ($\phi = 0.5$)

T	ρ	Model 1 ($\mu = 0, \delta = 0$)		Model 2 ($\mu = 1, \delta = 0$)		Model 3 ($\mu = 1, \delta = 1$)	
		KPSSM	S_nM	KPSSM	S_nM	KPSSM	S_nM

50	0.1	0.021	0.005	0.026	0.010	0.039	0.024
	0.85	0.238	0.110	0.194	0.118	0.166	0.049
	0.9	0.407	0.216	0.324	0.226	0.276	0.080
	0.95	0.608	0.375	0.498	0.398	0.402	0.165
	0.97	0.697	0.462	0.592	0.480	0.457	0.199
	0.99	0.782	0.581	0.692	0.601	0.483	0.226
	1	0.812	0.652	0.756	0.672	0.486	0.236
100	0.1	0.028	0.005	0.029	0.009	0.032	0.004
	0.85	0.123	0.047	0.100	0.048	0.083	0.016
	0.9	0.290	0.181	0.260	0.184	0.228	0.077
	0.95	0.596	0.464	0.539	0.471	0.481	0.282
	0.97	0.737	0.590	0.669	0.603	0.598	0.393
	0.99	0.853	0.739	0.802	0.750	0.673	0.487
	1	0.893	0.813	0.849	0.819	0.687	0.499
500	0.1	0.039	0.018	0.035	0.022	0.029	0.012
	0.85	0.044	0.013	0.036	0.018	0.034	0.004
	0.9	0.047	0.012	0.043	0.016	0.034	0.004
	0.95	0.281	0.257	0.291	0.257	0.344	0.214
	0.97	0.571	0.621	0.623	0.623	0.716	0.601
	0.99	0.925	0.916	0.923	0.918	0.951	0.910
	1	0.983	0.987	0.987	0.987	0.969	0.946

Table 4: Size and power, MA(1) errors ($\theta = 0.5$)

T	ρ	Model 1 ($\mu = 0, \delta = 0$)		Model 2 ($\mu = 1, \delta = 0$)		Model 3 ($\mu = 1, \delta = 1$)	
		KPSSM	S_nM	KPSSM	S_nM	KPSSM	S_nM
50	0.1	0.017	0.033	0.035	0.045	0.097	0.106
	0.85	0.093	0.013	0.115	0.035	0.162	0.022
	0.9	0.076	0.009	0.117	0.029	0.188	0.018
	0.95	0.075	0.005	0.121	0.027	0.197	0.016
	0.97	0.116	0.004	0.130	0.029	0.194	0.016
	0.99	0.218	0.009	0.165	0.033	0.189	0.016
	1	0.295	0.020	0.214	0.039	0.193	0.017
100	0.1	0.017	0.019	0.019	0.022	0.028	0.045
	0.85	0.101	0.047	0.134	0.083	0.188	0.062
	0.9	0.083	0.035	0.115	0.061	0.178	0.045
	0.95	0.054	0.007	0.078	0.036	0.155	0.037
	0.97	0.111	0.005	0.081	0.029	0.150	0.030
	0.99	0.288	0.016	0.141	0.035	0.156	0.029
	1	0.411	0.068	0.264	0.081	0.145	0.022
500	0.1	0.029	0.018	0.024	0.020	0.020	0.015
	0.85	0.089	0.056	0.088	0.065	0.104	0.059
	0.9	0.082	0.038	0.087	0.051	0.100	0.040

0.95	0.066	0.014	0.073	0.028	0.074	0.014
0.97	0.050	0.003	0.048	0.009	0.058	0.001
0.99	0.287	0.087	0.175	0.088	0.112	0.007
1	0.741	0.510	0.634	0.513	0.281	0.070

Examining the results presented in these three tables, some points are worth noting.

First, the KPSSM has a higher probability of rejection of the null hypothesis than the S_nM test. This difference decreases when the sample size is larger, $T = 500$. This implies that when the sample size is small, $T = 50$ or 100 , the KPSSM presents higher power but also higher empirical sizes than the S_nM test.

Second, the presence of deterministic terms in the DGP leads to smaller power values in both tests, as has been observed in other tests in the literature.

Third, when the disturbance follows an AR(1) process, the probability of rejection of the null hypothesis of both tests is higher than in the case of a white noise. However, when the disturbance follows a MA(1) process, the probability of rejection is smaller in all cases, even with a large sample, and the presence of a linear trend has a more perverse effect than when the disturbance is either a white noise or follows an AR(1) process.

Since, in all the situations, the KPSSM test exhibits higher power but also higher empirical size than the S_nM test, it is not clear which test has better small sample properties. So, in the following experiments, as in [2, 4, 8], we compare the tests by determining critical values⁶ for each, such that the rejection rates coincide at some prespecified value (50% in our experiment) under the

⁶ The critical values for the KPSSM and S_nM tests, for a given power of 50%, have been calculated using 50,000 replications.

alternative hypothesis (power) and, then, we compare their size. So, one test dominates another if it consistently has a smaller size. We do this because the null hypothesis is not simple. The results are illustrated in Figures 1, 2 and 3.

When the disturbance is a white noise -Figure 1- it can be seen that, when the sample size is small ($T = 50$), the KPSSM test outperforms the S_nM test when the model has no deterministic terms (Model 1) or a trend term (Model 3). When the sample size is larger ($T = 100, 500$), both tests perform similarly.

Similar results are obtained when the disturbance follows an $AR(1)$ process with $\phi = 0.5$, as can be seen in Figure 2.

However, the results change when the disturbance follows a $MA(1)$ model with $\theta = 0.5$. As can be seen in Figure 3, the KPSSM test outperforms the S_nM test, especially for Models 1 and 2, as well as for Model 3 with a large sample size.

In general, the KPSSM test shows a better performance, in terms of the trade-off between size and power, than the S_nM , and this better performance is more visible when the disturbance follows a $MA(1)$ process.

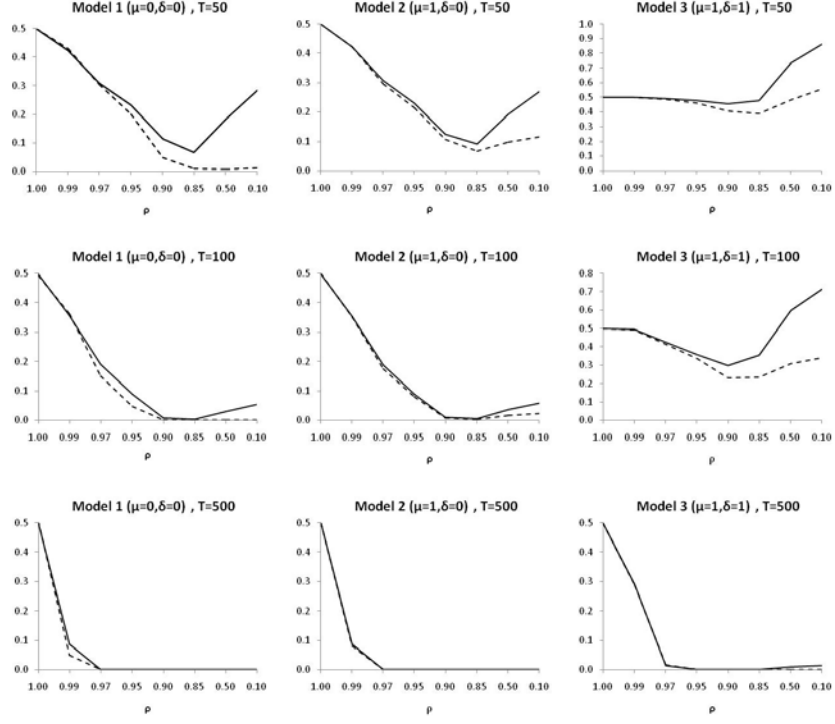


Figure 1: Size of the tests when the power is 0.5 and iid errors: - - - KPSSM, — S_nM

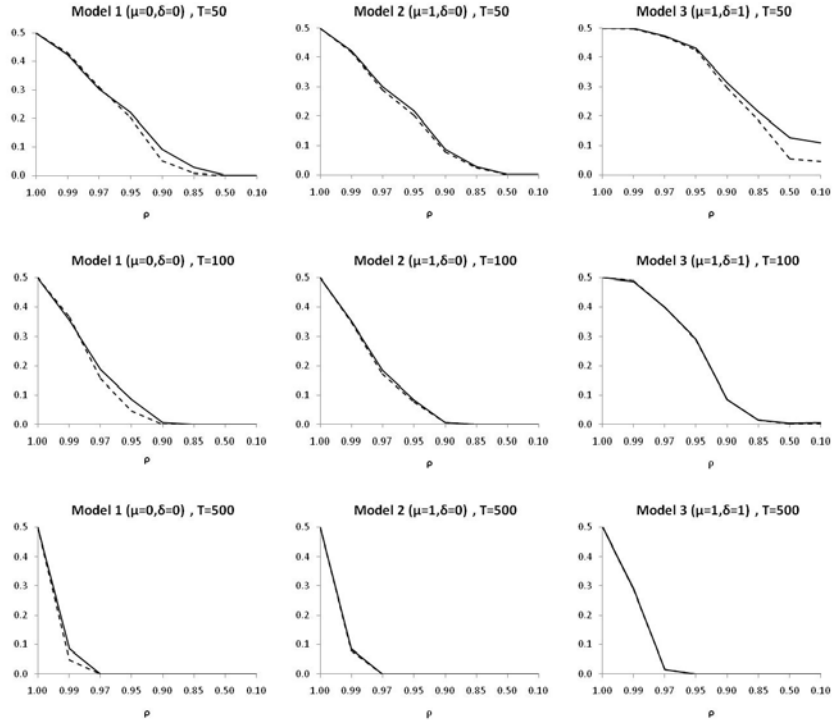


Figure 2: Size of the stationarity tests when the power is 0.5 and AR(1) errors: - - - KPSSM, — S_nM

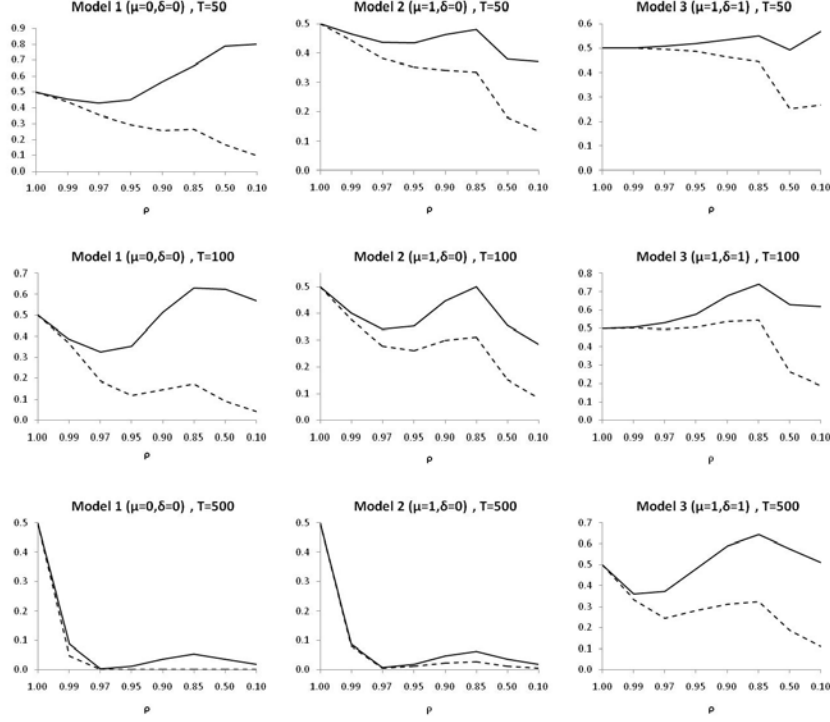


Figure 3: Size of the stationarity tests when the power is 0.5 and MA(1) errors: - - - KPSSM, — S_nM

4. Conclusions

In this paper, we have compared two procedures to test for the null hypothesis of stationarity against the alternative of the presence of a unit root, namely, the KPSS test proposed in [6] and the S_n test developed in [13]. In both cases, we have used the estimator of the long-run variance with the boundary condition rule proposed in [12] because, in a first Monte Carlo study, we concluded that this estimator improves the finite sample properties of the original test of Xiao. In general, the conclusion of the second experiment is that the modified version of the KPSS test shows a better performance, in terms of the trade-off between size and power, than that of the modified Xiao test. Furthermore, this better performance is more visible when the disturbance follows a MA(1) process than when the disturbance is either a white noise or follows an AR(1) process. So, in applied research, when we have to test whether a series is stationary, we recommend using the KPSSM test rather than the S_nM test.

Acknowledgements

We thank the Associate Editor and three anonymous referees for their very helpful comments. This study has benefited from the financial support of the research projects ECO 2009-07936, ECO2015-65582 and ECO2016-74940-P financed by the MEC-MICINN and the consolidated research group of ‘Selección de modelos econométricos’ (DGA-S21) financed by the Department of Science, Technology and Universities of the Aragonese Government and the European Regional Development Fund.

References

- [1] D.W.K. Andrews, Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econom.* 59 (1991) 817-858.
- [2] M.I. Ayuda, A. Aznar, A Monte Carlo study to compare two recent modifications of the KPSS test for near integration. *Appl. Econ. Lett.* 18 (2010) 1759-1764.
- [3] J.L. Carrion-i-Silvestre, A. Sansó, A guide to the computation of stationarity tests, *Empir. Econ.* 31 (2006) 433-448.
- [4] D. Harris, S. Leybourne, B. McCabe, Modified KPSS tests for near integration. *Econom. Theory* 23 (2007) 355-363.
- [5] B. Hobijn, P. H. Franses, M. Ooms, Generalizations of the KPSS-test for stationarity, *Stat. Néerl.* 58 (2004) 483-502.
- [6] D. Kwiatkowski, P.C.B. Phillips, P. Schmidt, Y. Shin, Testing the null hypothesis of stationarity against the alternative of a unit root, *J. Econom.* 54 (1992) 159-178.
- [7] S.J. Leybourne, B.P.M. McCabe, A consistent test for a unit root, *J. Bus. Econ. Stat.* 12 (1994) 157-166.
- [8] U. Müller, Size and power of tests of stationarity in highly autocorrelated time series. *J. Econom.* 128 (2005) 195-213.
- [9] S. Nabeya, K. Tanaka, Asymptotic theory of a test for constancy of regression coefficients against the random walk alternative. *Ann. Stat.* 16 (1988) 218-35.

- [10] P. Saikkonen, R. Luukkonen, Testing for a moving average unit root in Autoregressive Integrated Moving Average Models. *J. Am. Stat. Assoc.* 88 (1993) 596-601.
- [11] G.E. Schwarz, Estimating the dimension of a model. *Ann. Stat.*, 6 (1978) 461-464.
- [12] D. Sul, P.C.B. Phillips, C.Y. Choi, Prewhitening bias in HAC estimation, *Oxf. Bull. Econ. Stat.* 67 (2005) 517-546.
- [13] Z. Xiao, Testing the null hypothesis of stationarity against an autoregressive unit root alternative. *J. Time Ser. Anal.* 22 (2001) 87-105.