JETEM

Journal of Economic Theory and Econometrics

The Journal of the Korean Econometrics Society

(계 량 경 제 학 보)

March 2009

Volume 20. Issue 1. pp.1-14

Improving empirical size of the KPSS test of stationarity*

InChoi†

Abstract This note proposes a new testing procedure that can alleviate the size prob-Lem associated with semiparametric tests of stationarity. The note is focused on The test by Kwiatkowski, Phillips, Schmidt and Shin (1992) considering its pop-ularity in the literature. The testing procedure of this note employs sample-split and the Bonferroni test. The sample is split into two parts, one corresponding to the odd index and the other to the even index, and the KPSS test is applied to each subsample. The two KPSS tests are then combined using the Bonfer-roni principle. Simulation results demonstrate that this procedure significantly reduces the size distortion of the KPSS test.

Keywords stationarity test, size distortion, sample-split, Bonferroni proce-dure

JEL Classification C32

* I thank Lutz Kilian for his comments on this note.

† Department of Economics, Sogang University, Shinsu-dong 1, Mapo-gu, Seoul, Korea. E-mail: inchoi@sogang.ac.kr.

Received April 3, 2008, Revised July 29, 2008, Accepted August 5, 2008

1. Introduction

Along with unit root tests, tests for the null of stationarity have often been used in practice. Among these, tests by Kwiatkowski, Phillips, Schmidt and Shin (1992; $KPSS$ hereafter) and Leybourne and McCabe (1994) appear to be most popular in the empirical literature. Other tests include Saikkonen and Luukkonen (1989), Tanaka (1990), Bierens and Guo (1993), Choi (1994) and Müller (2005)[1], to name a few.

Though the $KPSS$ test has been widely used, Caner and Kilian (2001) show by simulation that it is subject to immense size distortions when the null is close to the alternative of a unit root. The purpose of this note is to propose a new testing procedure that alleviates the size problem of the $KPSS$ test. Lanne and Saikkonen (2003) also suggest methods that improve size properties of the parametric tests of stationarity such as Saikkonen and Luukkonen (1989) and Leybourne and McCabe (1994). The testing procedure of this note can also be applied to other semiparametric tests of stationarity such as Tanaka (1990) and Choi (1994). However, we will focus on the $KPSS$ test considering its popularity in the literature.

The testing procedure of this note employs sample-split and the Bonferroni test. The sample is split into two parts, one corresponding to the odd index and the other to the even index, and the $KPSS$ test is applied to each subsample. The two $KPSS$ tests are then combined using the Bonferroni principle. Simulation results demonstrate that this procedure significantly reduces the size distortion of the $KPSS$ test. Sample-split has been used for

[1] However, unlike other tests, Müller tests the null of near unit roots against the alternative of a unit root. The relative merits of this and other existing tests in finite samples have not been studied yet.

many different purposes in econometrics. The reader is referred to Dufour and Torrès (1998) for other applications.

This note is organized as follows. Section 2 introduces the $KPSS$ test using the sample-split and the Bonferroni procedure. Section 3 reports simulation results.

2. $KPSS$ test using sample-split

Consider the model

$$y_t = \mu + \rho y_{t-1} + u_t,\ t = 1, ..., T, \tag{1}$$

where u_t is an $I(0)$ time series and y_0 is assumed to be a fixed constant. For later convenience, assume that T is an even number. We are interested in the null hypothesis of level stationarity,

$$H_0 : |\rho| < 1,$$

against the alternative hypothesis of a unit root,

$$H_1 : \rho = 1 \text{ and } \mu = 0.$$

Let the residual from regressing y_t on 1 and y_{t-1} be $\hat{u}_t$. The $KPSS$ test statistic for the null of level stationarity is

$$KPSS = \frac{1}{\hat{\sigma}_u^2 T} \sum_{t=1}^{n} s_t^2,$$

where $s_t = \sum_{i=1}^{t} \hat{u}_i$ and $\hat{\sigma}_u^2$ is is the long-run variance estimator using $\{\hat{u}_t\}$. The limiting distribution of this test statistic is $\int_0^1 V_1(r)^2 dr$, where $V_1(r) = W(r) - rW(1)$ is a standard Brownian bridge.

When ρ is close to 1, the $KPSS$ test is known to suffer from size distortions. In order to make the value of the AR(1) coefficient less close to 1, consider the sample-split

$$\begin{aligned} y_{2t} &= \mu + \rho y_{2t-1} + u_{2t} \\ &= (\mu + \rho\mu) + \rho^2 y_{2t-2} + u_{2t} + \rho u_{2t-1} \\ &= \mu' + \rho^2 y_{2t-2} + v_{2t}, \text{ say,} \end{aligned}$$

and

$$y_{2t-1} = \mu' + \rho^2 y_{2t-3} + w_{2t-1}, \text{ say,}$$

where $\mu' = \mu(1+\rho)$, $v_{2t} = u_{2t} + \rho u_{2t-1}$ and $w_{2t-1} = u_{2t-1} + \rho u_{2t-2}$. The value of the AR(1) coefficient for these two subsamples, $\{y_{2t}\}_{t=1}^{T/2}$ and $\{y_{2t-1}\}_{t=1}^{T/2}$, is less close to 1 than that for the original sample. Thus, using one of these two subsamples is expected to yield better empirical size for the $KPSS$ test. However, discarding half of the sample in order to choose one of the two will certainly involve power loss. So we will consider combining the two tests.

To this end, let the $KPSS$ test statistics using $\{y_{2t}\}_{t=1}^{T/2}$ and $\{y_{2t-1}\}_{t=1}^{T/2}$ be $KPSS_e$ and $KPSS_o$, respectively, and consider

$$SS\text{-}KPSS = \max(KPSS_e, KPSS_o).$$

We will call $SS\text{-}KPSS$ sample-split $KPSS$ test statistic. Letting $c_{\alpha/2}$ be the $\frac{\alpha}{2}$-level critical value from the distribution $\int_0^1 V_1(r)^2 dr$,[2] we obtain by using

[2] That is, $P[\int_0^1 V_1(r)^2 dr \geq c_{\alpha/2}] = \frac{\alpha}{2}$.

the Bonferroni inequality,

$$\begin{aligned}
\lim_{T\to\infty} P[SS\text{-}KPSS \ \leq \ c_{\alpha/2}] &= \lim_{T\to\infty} P[KPSS_e \leq c_{\alpha/2} \text{ and } KPSS_o \leq c_{\alpha/2}] \\
&= 1 - \lim_{T\to\infty} P[KPSS_e > c_{\alpha/2} \text{ or } KPSS_o > c_{\alpha/2}] \\
&\geq 1 - \left(\lim_{T\to\infty} P[KPSS_e > c_{\alpha/2}] + \lim_{T\to\infty} P[KPSS_o > c_{\alpha/2}] \right) \\
&= 1 - \alpha
\end{aligned}$$

or equivalently

$$\lim_{T\to\infty} P[SS\text{-}KPSS \ \geq c_{\alpha/2}] \leq \alpha.$$

This inequality implies that the SS-$KPSS$ test using $\frac{\alpha}{2}$-level critical values from the distribution of $\int_0^1 V_1(r)^2 dr$ has an asymptotic size less than or equal to α. In other words, using the $\frac{\alpha}{2}$-level critical values of the $KPSS$ test for the SS-$KPSS$ test yields the asymptotic size less than or equal to α.

The testing procedure using the SS-$KPSS$ test statistic is expected to alleviate the size distortion problem of the $KPSS$ test for two reasons. First, as mentioned previously, the subsamples we use are based on an AR model with coefficients less close to 1 than that of the original model. Second, the Bonferroni procedure makes the test more conservative. However, the Bonferroni procedure itself does not alone improve the $KPSS$ test. If the original sample is cut half and the Bonferroni procedure is applied to the $KPSS$ tests from the two samples, no improvement of the $KPSS$ test is observed according to simulation results not reported in this note.

The same procedure can be used to test the null of trend-stationarity. For this null hypothesis, model

$$y_t = \mu + \beta t + \rho y_{t-1} + u_t, \ t = 1, ..., T$$

is used instead of model (1), and residual $\hat{u}_t$ is obtained from regressing y_t on 1, t and y_{t-1}. The limiting distribution of the $KPSS$ test is $\int_0^1 V_2(r)^2 dr$, where $V_2(r) = W(r) + (2r - 3r^2)W(1) + (-6r + 6r^2)\int_0^1 W(s)ds$. The sample-split $KPSS$ test is formulated in the same way using the new residuals.

The sample-split considered so far separates the whole sample into two pieces. We may go on further and split the sample into more than two pieces. But this will make the SS-$KPSS$ test more conservative. According to unreported simulation results, using three pieces makes empirical power close to zero at the 10% level when $T = 100$ and $y_0 = 0$, though the empirical power improves with increasing sample sizes. So if one wants be very conservative about rejecting the null, a large number of sample splitting serves the purpose and vice versa.

3. Simulation

This section reports simulation results for the SS-$KPSS$ and $KPSS$ tests. In order to study the finite-sample properties of the SS-$KPSS$ and $KPSS$ tests for the null of level stationarity, data were generated by

$$\begin{aligned} y_t &= \mu + \rho y_{t-1} + u_t, \ (t = 1, ..., T) \\ u_t &\sim iid\ N(0, \sigma^2). \end{aligned}$$

Since the parameters μ and σ^2 do not affect the size and power of the SS-$KPSS$ and $KPSS$ tests, we set $\mu = 0$ and $\sigma^2 = 1$. Values of ρ that we considered are 0, 0.5, 0.7, 0.8, 0.9, 0.95, 0.98, 0.99 and 1. The size of the tests depends on the initial value y_0, but not their power. We considered 0, 5 and 10 for the value of y_0. The usual scheme of discarding the first few observations amounts to using an initial value very close to 0 under the null.[3] The number of iterations was 20,000. For the long-run variance estimation, we used $[12(sample\ size/100)^{0.25}]$ for the lag length suggested in $KPSS$ and the Bartlett window. Using another choice of the lag length in $KPSS$, $[4(sample\ size/100)^{0.25}]$, gave less desirable results. It is well known that the kernel window for the long-run variance estimation (for example, see Newey and West, 1994) is less important than the lag length. Experimental results for the null of level stationarity are reported in Table 1. Note that the empirical power for $y_0 = 5$ and $y_0 = 10$ are omitted in the table, since this is the same as that for $y_0 = 0$.

Experiments for the finite-sample properties of the SS-$KPSS$ and $KPSS$ tests for the null of trend-stationarity were performed in the same way except that the regression that generates residuals has t as an additional regressor. Note that the value of the time trend coefficient does not affect the tests. The relevant experimental results are reported in Table 2.

Tables 1 and 2 can be summarized as follows.

- The $KPSS$ test shows serious size distortions when ρ is close to 1 even at $T = 600$ for all the initial values. This corresponds with the results in Caner and Kilian (2001). The SS-$KPSS$ test significantly improves the size distortion problem of the $KPSS$ test in the vicinity of the alternative hypothesis. For example, when $T = 100, y_0 = 0$ and $\rho = 0.99$ in Table 1, the empirical size of the SS-$KPSS$ test is 3.56% at the 5% nominal level while that of the $KPSS$ test is 49.3% at the same nominal size. In almost all the entries in Tables 1 and 2, the empirical size improves when the SS-$KPSS$ test is used.

[3] The simulation results in Caner and Kilian (2001) correspond to this case.

- Empirical size for the null of trend-stationarity is closer to the nominal size or nominal level as shown in Table 2. But still, the problem of size distortions is serious for the $KPSS$ test and the SS-$KPSS$ test tends to alleviate it.

- Larger initial values bring more size distortions for both the $KPSS$ and SS-$KPSS$ tests at $T = 100,$ though their impact diminishes as the sample size grows.

- Though the SS-$KPSS$ test improves empirical size, empirical power decreases when it is used. In addition, it rejects too infrequently when the null is away from the alternative and $T = 100$ (see the case $\rho = 0$ at $T = 100$, for example). These are the price to be paid for improved empirical size in the vicinity of the alternative hypothesis.

- At $T = 100$ in Table 2, the SS-$KPSS$ test is very conservative at the 5% nominal level while the $KPSS$ test rejects too often in the corresponding case. Neither test appears to be satisfactory in this case.

4. Conclusion

We have proposed a new testing procedure that alleviates the size problem of the $KPSS$ test. Simulation results indicate the testing procedure works better than the $KPSS$ test when the null is close to the alternative. When to use the SS-$KPSS$ test instead of $KPSS$ depends on each researcher's choice. If he has a good reason to be conservative about rejecting the null of stationarity, he will prefer using the SS-$KPSS$ test.

References

Bierens, H. and S. Guo, 1993, Testing stationarity and trend stationarity against the unit root hypothesis, Econometric Reviews 12, 1–32.

Caner, M. and L. Kilian, 2001, Size distortions of the null hypothesis of stationarity: evidence and implications for the PPP debate, Journal of International Money and Finance 20, 639–657.

Choi, I., 1994, Residual based tests for the null of stationarity with applications to U.S. macroeconomic time series, Econometric Theory 19, 720–746.

Dufour, J.-M. and O. Torrès, 1998, Union-intersection and sample-split methods in econometrics with applications to SURE and MA models, in Handbook of Applied Economic Statistics (eds. D. Giles and A. Ullah), Chapter 14, 465–505 (Marcel Dekker, New York).

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt and Y. Shin, 1992, Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root?, Journal of Econometrics 54, 159–178.

Lanne, M. and P. Saikkonen, 2003, Reducing size distortions of parametric stationarity tests, Journal of Time Series Analysis 24, 423–439.

Leybourne, S.J. and McCabe, B.P.M., 1994, A consistent test for a unit root, Journal of Business and Economic Statistics 12, 157–166.

Müller, U., 2005, Size and power of tests of stationarity in highly autocorrelated time series, Journal of Econometrics, vol. 128, no. 2, pp. 195-213.

Newey and West, 1994, Automatic lag selection in covariance matrix estimation, Review of Economic Studies 61, 631–653.

Tanaka, K., 1990, Testing for a moving average unit root, Econometric Theory 6, 411–432.

Saikkonen, P. and R. Luukkonen, 1989, Testing for a moving average unit root in autoregressive integrated moving average models, Journal of the American Statistical Association 88, 596–601.

Table 1: Empirical size and power of the SS-$KPSS$ and $KPSS$ tests: Intercept only

		$T = 100$				$T = 300$				$T = 600$			
		SS-KPSS		*KPSS*		*SS-KPSS*		*KPSS*		*SS-KPSS*		*KPSS*	
	ρ	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
$y_0 = 0$	**0.0**	0.08	3.63	3.39	9.93	2.34	7.22	4.38	9.60	3.85	9.02	4.65	10.1
	0.5	0.08	3.99	4.60	12.4	2.47	7.07	5.43	11.6	3.51	7.84	5.80	11.9
	0.7	0.15	4.48	6.95	16.3	2.82	7.24	7.38	14.6	3.43	7.44	7.33	14.1
	0.8	0.10	4.61	9.34	20.1	2.90	7.82	10.1	18.6	3.83	8.30	8.95	16.8
	0.9	0.20	6.97	18.0	31.8	5.45	11.6	18.6	29.5	6.04	11.7	16.7	26.8
	0.95	0.46	11.8	28.5	43.2	11.1	20.3	34.1	47.8	12.3	20.1	32.5	45.5
	0.98	1.29	20.5	41.4	56.1	25.4	37.0	55.4	69.4	30.0	40.3	59.7	72.8
	0.99	3.56	27.9	49.3	62.4	36.5	48.5	66.4	79.0	45.8	56.6	75.8	85.4
	1.0	9.62	42.5	60.7	71.4	60.4	68.7	82.6	89.9	76.3	83.0	92.5	96.1
$y_0 = 5$	**0.0**	0.37	7.25	6.42	15.8	3.43	9.58	5.47	11.7	4.11	9.65	5.05	10.9
	0.5	0.72	8.46	8.46	19.0	3.54	9.03	7.23	14.4	4.06	8.73	6.45	12.8
	0.7	0.94	9.73	11.8	23.2	3.69	9.31	9.12	16.8	3.90	8.17	7.74	15.0
	0.8	1.25	11.3	16.1	28.6	4.60	9.81	11.8	20.9	4.15	8.81	9.83	17.4
	0.9	1.10	16.6	27.9	41.8	6.92	14.0	22.0	33.7	7.13	12.5	17.7	28.7
	0.95	2.93	23.9	41.3	54.8	14.8	24.4	38.1	52.0	13.6	21.4	34.6	47.5
	0.98	3.82	28.8	48.4	61.2	30.1	41.7	59.3	72.9	31.8	42.5	61.9	74.4
	0.99	4.67	31.6	52.4	65.1	40.1	51.3	69.4	81.3	47.5	58.3	76.3	85.6
$y_0 = 10$	**0.0**	0.24	7.87	8.65	23.0	4.52	12.9	7.28	15.3	5.00	11.5	6.61	13.0
	0.5	0.76	11.4	13.2	29.6	4.85	11.9	9.64	18.6	4.98	10.5	8.19	15.5
	0.7	1.23	13.7	18.4	35.9	5.36	12.1	12.9	23.0	5.05	10.5	10.0	18.0
	0.8	1.90	16.7	25.6	44.0	6.50	13.8	16.4	27.9	5.73	11.2	12.6	21.7
	0.9	4.53	28.5	45.0	62.5	10.8	19.8	29.4	42.4	9.06	15.7	22.0	33.2
	0.95	9.56	44.2	63.0	75.3	21.3	32.4	47.3	61.8	17.2	26.1	38.5	52.4
	0.98	9.86	48.1	65.7	75.9	41.6	52.8	69.1	80.5	37.8	48.2	67.9	79.5
	0.99	7.94	41.0	60.0	70.6	50.3	60.6	75.6	85.3	53.4	64.0	80.8	89.2

Table 2: Empirical size and power of the SS-$KPSS$ and $KPSS$ tests: Intercept and linear time trend

		$T = 100$				$T = 300$				$T = 600$			
		SS-KPSS		*KPSS*		*SS-KPSS*		*KPSS*		*SS-KPSS*		*KPSS*	
	ρ	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
$y_0 = 0$	**0.0**	0.04	6.10	3.28	11.2	2.19	7.48	4.25	10.0	3.11	7.71	4.48	10.1
	0.5	0.08	5.78	4.93	14.4	2.28	6.96	5.82	12.7	3.00	7.44	5.71	12.3
	0.7	0.10	5.57	6.76	18.1	2.43	7.24	7.93	16.4	3.47	8.16	7.72	15.3
	0.8	0.06	5.59	9.60	23.1	2.91	7.80	11.4	21.6	3.93	8.41	10.8	20.1
	0.9	0.05	7.26	18.9	36.7	5.41	12.3	24.4	38.9	6.95	13.5	22.4	35.4
	0.95	0.02	11.9	29.7	48.9	12.5	23.2	45.2	61.4	15.3	25.4	45.5	60.9
	0.98	0.21	18.3	39.3	58.2	27.7	42.3	69.0	81.2	40.1	53.9	76.8	86.9
	0.99	0.18	20.7	42.7	61.1	37.3	52.8	77.2	86.5	57.5	69.8	87.9	93.8
	1.0	0.18	21.8	43.7	61.9	45.7	59.9	81.9	89.8	73.5	82.8	94.1	97.2
$y_0 = 5$	**0.0**	0.35	15.1	8.04	20.7	3.37	10.2	5.54	12.5	4.22	9.70	5.17	11.2
	0.5	0.54	14.9	10.5	24.4	3.45	9.60	7.96	16.2	4.15	8.86	6.81	13.8
	0.7	0.54	14.2	13.6	28.7	3.88	9.62	10.1	19.6	4.11	8.73	8.94	17.2
	0.8	0.37	14.5	18.3	35.2	4.26	10.7	14.3	25.9	4.47	9.62	12.2	22.0
	0.9	0.26	15.8	28.5	47.0	7.66	15.5	28.3	43.8	7.83	14.9	24.0	37.4
	0.95	0.14	17.1	35.4	53.6	15.3	27.4	50.4	65.8	17.1	27.6	47.9	63.7
	0.98	0.18	19.0	40.4	59.3	30.5	45.2	70.9	82.2	42.1	55.8	78.8	88.3
	0.99	0.17	20.7	43.3	61.4	38.4	53.6	78.0	87.2	58.8	71.1	88.5	94.1
$y_0 = 10$	**0.0**	0.31	21.0	13.1	35.5	4.71	14.8	8.62	18.8	5.61	13.0	6.93	14.1
	0.5	0.72	24.2	18.7	42.3	5.38	14.2	11.9	23.0	5.71	12.1	9.76	18.0
	0.7	1.03	24.8	25.2	48.4	6.49	15.0	15.7	29.2	5.57	11.7	11.6	21.6
	0.8	1.08	26.2	33.3	56.1	7.57	17.0	21.8	36.2	6.32	12.9	15.9	27.0
	0.9	0.96	31.6	48.0	67.6	13.0	24.1	39.0	55.4	11.1	19.8	30.2	44.6
	0.95	0.55	30.4	49.6	66.3	23.8	37.4	59.9	74.2	23.0	34.8	55.0	70.0
	0.98	0.16	21.5	43.7	61.3	37.2	51.6	75.3	85.4	47.6	60.8	82.1	91.0
	0.99	0.13	21.2	43.4	61.5	40.5	55.1	78.3	87.7	61.5	72.9	89.3	94.7