行政院國家科學委員會專題研究計畫　成果報告

t 檢定偵測結構性變動的穩健性

計畫類別：個別型計畫
計畫編號：NSC91-2415-H-004-001-
執行期間：91 年 08 月 01 日至 93 年 07 月 31 日
執行單位：國立政治大學國際貿易學系

計畫主持人：郭炳伸

報告類型：精簡報告
報告附件：國外研究心得報告
處理方式：本計畫可公開查詢

中　華　民　國 93 年 11 月 1 日

Bootstrap Inference for Stationarity [*]

Biing-Shen Kuo [†]
National Chengchi University

Abstract

Tests for the stationarity null due to Kwiatkowski et al. (1992) continue to be an indispensable part of tool kits for empirical researchers when investigating time series property of aggregate variables. As well-documented in the literature (see for instance, Caner and Kilian, 2001), the tests display considerable size distortions, if the data generated under the null is highly persistent. The paper offers an asymptotic explanation in a local-to-unity framework. Our analytical derivations unveil that the tests fail to converge without a re-normalization. The surprising finding suggests that the size bias deteriorates as sample size increases, but declines as bandwidth number increases, consistent with simulation evidence. The derivations however give little clue to how to mitigate the size bias, because of an inability to consistently estimate the local-to-unity parameter. While it is natural to appeal to the bootstrapping, it proves infeasible to construct a sensible re-sampling scheme, based on the unobserved component model from which the observed series is generated. We resolve the difficulty by drawing bootstrap samples from a parametric ARIMA model, second-order equivalent in moments to the unobserved component model. Even in the presence of highly persistent processes, our bootstrap tests are found to yield very satisfactory control over the rejection probability at little cost of power loss.

JEL classification: C12, C14, C15, C22
Keywords: bootstrap, ARMA model

[*]The author gratefully acknowledges grants from the National Science Council.

[†]Corresponding author, Graduate Institute of International Trade, National Chengchi University, Taipei 116, Taiwan; Tel: (886 2) 2939 3091 ext. 81029, Fax: (886 2) 2937-9071, Email: bsku@nccu.edu.tw

1 Introduction and Summary

Tests for the stationarity null has appeared to be an indispensable part of tool kits when investigating time series property of aggregate variables. Information about the persistence nature of the observed series coming from evidence with the tests often complements to that from existing unit root tests. On the other hand, the spirit of testing for the stationarity null may be more consistent with classical hypothesis testing where the hypothesis to be tested under the null is the one that researchers believe in when testing for some economic theories. For instance, many international macroeconomists tend to hold the view that relative price levels between countries display at most transitory deviations from purchasing power parity, as a result of market forces. The hypothesis that real exchange rates are mean-reverting is thus natural to be tested under the null, and should not be rejected lightly unless strong evidence against it is established. A partial list for available stationarity tests that possess these features could consist of Kwiatkowski, Phillips, Schmidt and Shin (1992) (short for KPSS, hereafter), Saikkonen and Luukkonen (1993), and Leybourne and McCabe (1994).

For inference on stationarity to be able to be drawn reliably from empirical evidence, test statistics on which statistical decisions are based at least ought to demonstrate robust control over the rejection frequencies. While a minimal requirement for the test statistics, satisfactory size control has proved very difficult to meet when the series under test are stationary but highly persistent processes. Caner and Kilian (2001) offered a comprehensive account of the size problem with stationarity tests in the context. Simulations by KPSS (1992) already revealed potential size distortions about their tests. Specifically, there are considerable over-rejections when the simulated data is drawn from simple autoregressive models with the persistence parameter closer to unit root, for sample sizes that mostly encounter in practice. Moreover, an increase in samples does not help reduce but aggravate occurrence of rejections. The latter finding runs counter to the idea of large sample theory on which stationarity tests typically rely: asymptotic approximations yield more accuracy as sample increases. The existence of size problem immediately calls into questions the credibility of empirical evidence with stationarity tests. It is well understood that many observed time series in empirical macroceconomics and international finance often exhibit a strong persistence, and thus fall into the problematic parameter zone. In the presence of spurious rejections, a clear interpretation of rejections by the stationarity tests now turns out to be a formidable task, whether or not the true processes underlying the considered series are stationary.

The purposes of our paper are two fold. The first is to provide a theoretical underpinning for the sources of distorted sizes in stationarity tests. We concentrate on the KPSS tests that have been widely applied in empirical work, due to a much less computation involving in correcting for error autocorrelation, comparing to the tests of Saikkonen and Luukkonen (1993), and Leybourne and McCabe (1994). The semiparametric correction is accomplished through an estimation of the "long-run variance" accounting for a wide ranges of short-run dynamics. In a way, the analysis carried out here is parallel to the development of the literature in the unit root testing. It is well known that conventional unit root tests, such as Phillips and Perron tests (Phillips and Perron, 1988) and their modified variants (Perron and Ng, 1996), subject to dramatic size distortions when the error process is close to the unit circle. Thus, the size problem with the KPSS tests shares a very similar nature as that with the aforementioned unit root tests, where the estimation of the long run variance plays an important role in shaping asymptotic behaviors of either classes of tests. Using the local-to-unity framework, developed by Phillips (1987) and Nabeya and Perron (1994), the simulation findings about the KPSS tests are able to be reconciled with our analytical results. There are two important messages emerging from our asymptotic analysis. First, in the presence of stationary but highly persistent process, the KPSS tests can not converge to sensible limit distributions without a re-normalization that depend on the ratio of sample size to bandwidth number, explaining the counter-intuitive simulations of Caner and Kilian (2001). Second, the growth rate of the bandwidth number has an impact on the null asymptotic of the KPSS tests. As long as it grows at a rate slower than sample sizes, increasing in sample sizes would never alleviate the size problem.

While our analytical results are useful in explaining why the KPSS tests suffer from size distortions, they do not lend practical solutions to resolving the problem. This is because the asymptotics obtained for the KPSS tests under the null in the local context depends on unknown local-to-unity coefficient that can not be consistently estimated. Therefore, our theory is indicative of the impossibility to reducing the size distortion using asymptotic argument.

Our second purpose is to develop a bootstrap procedure that can have actual finite sample rejection frequencies closer to asymptotic nominal levels. The development of such a bootstrap stationarity test does not come as straightforward as that of the bootstrap unit root tests. The major difficulty for doing so lies in a lack of a reasonable model for bootstrap samples to be generated from under the null of the KPSS tests. The component model on which the KPSS tests are based on basically renders little sampling structure under the null. We resolve the difficulty by drawing bootstrap samples from a parametric ARIMA model which is second-order equivalent in moments to the unobserved component model. Our simulations demonstrate that even in the presence of highly persistent processes, our bootstrap tests are found to yield very satisfactory control over the rejection probability at little cost of power loss. This suggests a substantial gains from applying our bootstrap stationarity tests in terms of an asymptotic refinement over the first oder asymptotics.

The rest of the paper is summarized as follows. Section 2 introduces the test statistics. Major theoretical results are given in Section 3 in a local-to-unity context. The idea and the detailed re-sampling procedure of our bootstrap stationarity test are taken up in Section 4.

2 Test Statistics

Tests for the stationarity null mounted by Kwiatkowski et al. (1992) is derived from a component model that consists of a deterministic component, a random walk and a stationary error:

$$y_t = \sum_{i=0}^{m} \beta_i t^i + r_t + \epsilon_t, \; t = 1 \ldots T, \tag{1}$$

where m could be either 0 or 1 that represents intercept or both intercept and deterministic time trend, respectively; and r_t is a random walk, in which

$$r_t = r_{t-1} + \zeta_t,$$

with fixed initial values r_0 set to zero without loss of generality, and ζ_t being independent stationary process. Whether the series under consideration y_t is stationary however hinges on the variance of random-walk error, σ_ζ^2. Given that ϵ_t is a stationary error, when $\sigma_\zeta^2 > 0$, y_t comes to be stationary only after differencing. Alternatively, the series is stationary around a constant level or a trend, if $\sigma_\zeta^2 = 0$. The hypothesis of interest thus can be formulated as

$$H_0 : \sigma_\zeta^2 = 0 \qquad \text{versus} \qquad H_1 : \sigma_\zeta^2 > 0 \tag{2}$$

The KPSS test is derived based on the Lagrange multiplier (LM) principle. The derivation of the test is equivalent to those considered by Nyblom (1986) and Nabeya and Tanaka (1988) to test for random coefficients. All these statistics are LBI tests and thus possess the optimal property that attains the highest power locally. The calculation of the LM-type test statistics is not as complicated as the derivation. First, regress y_t against an intercept (if $m = 0$), or an intercept and time trend (if $m = 1$), and obtain the residuals, denoted by $\hat{u}_t$. That is, $\hat{u}_t = y_t - \sum_{i=0}^{m} \hat{\beta}_i t^i$ in which $\hat{\beta}_i$ is OLS estimates of β_i. Next, compute the partial sum of the residuals, $S_t = \sum_{i=1}^{t} \hat{u}_i$, and estimate the long-run variance of ϵ_t, based on Newey and West (1987):

$$(3) \qquad \hat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T} \hat{u}_t^2 + 2\frac{1}{T}\sum_{i=1}^{L} w(i,L) \sum_{t=i+1}^{T} \hat{u}_t \hat{u}_{t-i}$$

where $w(i, L) = 1 - i/(1 + L)$ is Bartlett kernel, and L is bandwidth. Here to obtain a consistent estimate for the long-run variance, L needs to be increased as T increases. Our proofs of the main results are based on $L = o(T^{1/2})$ for the Bartlett kernel we employ. In simulations, $L = [k(T/100)]^{1/4}$, where k is constant, and $[\cdot]$ is the largest integer function. This follows from Schwert (1989). The LM test statistic can then be formed by

$$LM = T^{-2}\hat{\sigma}^{-2}\sum_{t=1}^{T} S_t^2$$

We shall denote the test statistic by $KPSS_\mu(k)$, and $KPSS_\tau(k)$, respectively, given $m = 0$ or 1.

Kwiatkowski et al. (1992) established that under the null and some regularity conditions, the limiting representations of $KPSS_\mu$ and $KPSS_\tau$ can be characterized as:

$$(4) \qquad KPSS_\mu \Rightarrow \int_0^1 V_\mu^2(r)dr, \quad KPSS_\tau \Rightarrow \int_0^1 V_\tau^2(r)dr$$

where $\Rightarrow$ denotes weak convergence, $V_\mu(r) = W(r) - rW(1)$ is a standard Brownian bridge, $V_\tau(r) = W(r) + (2r - 3r^2)W(1) + (-6r + 6r^2)\int_0^1 W(s)ds$, and $W(r)$ is a Wiener process. The tests reject the stationarity null for large values of the statistics by construction. Because these distributions are not standard and free of nuisance parameters, critical values at conventional significance levels needs to be computed via simulations, before the tests can have practical uses.

3 Local Asymptotic Behavior of the Tests

The usefulness of the limiting theory for the test statistics depends on if they can yield accurate approximations to the finite-sample distribution. It is by now well-documented that the KPSS tests subject to considerable distortions in the presence of highly persistent but stationary processes. Caner and Kilian (2001) illustrated and re-affirmed the points by

Table 1: Empirical Size performance of $KPSS_\tau(k)$ (no trend)

T	α	$KPSS_\tau(4)$	$KPSS_\tau(8)$	$KPSS_\tau(12)$
300	0.99	0.976	0.882	0.767
	0.98	0.963	0.832	0.698
	0.94	0.846	0.570	0.393
	0.30	0.073	0.062	0.056
600	0.99	0.995	0.957	0.887
	0.98	0.986	0.901	0.774
	0.94	0.847	0.575	0.382
	0.30	0.071	0.056	0.053

Note:
1. The rejection frequency in each entry is calculated based on a DGP $y_t = \alpha y_{t-1} + e_t$, with $e_t \overset{iid}{\sim} N(0,1)$, using asymptotic critical value at 5% nominal level (.146) in 5000 replications.
2. The test statistic $KPSS_\tau(k)$ is, as defined in the text, calculated at bandwidth number set to k.

providing systematic investigations on the tests. Before them, Monte Carlo simulations in KPSS (1992) have revealed potential size problems of their tests. Lee (1996) focused on the effect of bandwidth selection on both size and power of the tests.

Simulations of the sort delivers immediate relevance to interpretations of empirical evidence with the tests. It is not uncommon that aggregate time series that have been most examined are found to be highly persistent.

To appropriately address the size problems, Table 1 replicates simulation results reported in Caner and Kilian (2001), following their setup.

Summaries of the simulations:

1. For the same sample sizes, the closer α is to 1, the larger the size distortion of $KPSS_\tau(k)$ is.

2. For fixed values of α, the larger the values of k are, the less $KPSS_\tau(k)$ displays size distortion.

3. As α is closer to 1, an increase in sample size does not help to reduce the degree of size distortion with $KPSS_\tau(k)$. Surprisingly, the size distortion worsens off as sample sizes increase.

The asymptotic theory provided in Kwiatkowski et al. (1992) does not appear to be able to the aforementioned simulations.

Definiion 1: y_t is generated as follows:

$$y_t = \sum_{i=0}^{m} \beta_i t^i + r_t + u_t \tag{5}$$

$$u_t = (1 + c/T)u_{t-1} + \epsilon_t \tag{6}$$

$$r_t = r_{t-1} + \zeta_t \tag{7}$$

$$\zeta_t = (\lambda/T)e_t \tag{8}$$

Theorem 1 y_t is generated as by Definition 1. Assuming $\lambda = 0$,

1. If $m = 0$,

$$(\frac{L}{T})KPSS_\mu(k) \implies \frac{\int_0^1(\int_0^r \bar{J}_c(s)ds)^2 dr}{\int_0^1 \bar{J}_c^2(r)dr} \tag{9}$$

2. If $m = 1$,

$$(\frac{L}{T})KPSS_\tau(k) \implies \frac{\int_0^1(\int_0^r \tilde{J}_c(s)ds)^2 dr}{\int_0^1 \tilde{J}_c^2(r)dr} \tag{10}$$

where $\bar{J}_c(r) = J_c(r) - \int_0^1 J_c(s)ds$, $\tilde{J}_c(r) = J_c(r) + (6r - 4)\int_0^1 J_c(s)ds + (6 - 12r)\int_0^1 sJ_c(s)ds$, $J_c(r) = \int_0^r e^{c(r-x)}dW(x)$, W is a standard Wiener process defined on $[0, 1]$.

Theorem 2 Again assume that y_t is generated as by Definition 1. Also assume that $u_t = \epsilon_t$ in (6). As $T \to \infty$,

1. If $m = 0$,

$$KPSS_\mu(k) \implies \int_0^1 (\lambda\sigma_e\sigma_\epsilon^{-1}\int_0^r \bar{W}_2(s)ds + W_1^*(r))^2 dr$$

2. If $m = 1$,

$$KPSS_\tau(k) \implies \int_0^1 (\lambda\sigma_e\sigma_\epsilon^{-1}\int_0^r \tilde{W}_2(s)ds + W_1^{**}(r))^2 dr$$

where $\bar{W}_2(s) = W_2(s) - \int_0^1 W_2(r)dr$, $\tilde{W}_2(s) = W_2(s) + (6s-4)\int_0^1 W_2(r)dr + (6-12s)\int_0^1 rW_2(r)dr$, $W_1^*(s) = W_1(s) - sW_1(1)$, $W_1^{**}(s) = W_1(s) + (2s - 3s^2)W_1(1) + (-6s + 6s^2)\int_0^1 W_1(r)dr$, and W_1 and W_2 are mutually independent standard Wiener processes.

This theorem investigates the power behavior of the tests in the case where σ_ζ^2 is near zero while stationary component of the model. Note first that both tests, $KPSS_\mu(k)$ and $KPSS_\tau(k)$, share the same asymptotics in the limit.

Theorem 3 Assume that y_t is generated as by Definition 1. As $T \to \infty$,

1. If $m = 0$,

$$(\frac{L}{T})KPSS_{\mu}(k) \implies \frac{\int_0^1(\int_0^r \bar{J}_c(s)ds)^2dr}{\int_0^1 \bar{J}_c^2(r)dr}$$

2. If $m = 1$,

$$(\frac{L}{T})KPSS_{\tau}(k) \implies \frac{\int_0^1(\int_0^r \tilde{J}_c(s)ds)^2dr}{\int_0^1 \tilde{J}_c^2(r)dr}$$

where $\bar{J}_c(r)$ and $\tilde{J}_c(r)$ have been defined in Theorem 1.

Theorem 3 is interesting where it examines the power performance of the tests in the case where the variance governing the non-stationary component is small, while stationary component is highly persistent. Both test statistics are exactly the same as those obtained in Theorem 1 where the non-stationary component disappears. It is evident from the result that the stationary component dominates the non-stationary counterpart in the limit, due to the local set-up in the random-walk variance. The results have an important implication for empirical studies on the testing for stationarity. The results show that the test statistics display a spurious rejection of the stationarity null, using the conventional critical values, in the presence of highly persistent processes in the stationary component.

4 Bootstrap Tests for the Stationarity Null

4.1 Ideas

To bootstrap the test, we need to be able to approximate the distribution under the null through re-sampling. The component model, however, does not lend itself to such a possibility. The component model on which the test bases has a reduced form. The reduced form has a representation of parametric ARMA(1,1) model in the differenced series.

$$\Delta y_t = \alpha \Delta y_{t-1} + (1 - \theta L)\eta_t, \tag{11}$$

where $\eta_t \overset{iid}{\sim} (0, \sigma_\eta^2), \sigma_\eta^2 = \sigma_\epsilon^2/\theta$, and $\theta = \frac{1}{2}\{\frac{\sigma_\zeta^2}{\sigma_\epsilon^2} + 2 - (\frac{\sigma_\zeta^4}{\sigma_\epsilon^4} + 4\frac{\sigma_\zeta^2}{\sigma_\epsilon^2})^{1/2}\}$. where $\sigma_\zeta^2/\sigma_\epsilon^2$ is the signal-to-noise ratio.

The null that $\sigma_\zeta^2 = 0$ in the component model amounts to $\theta = 1$ in the reduced form, an moving average unit root. This equivalence relation can be employed to construct the bootstrap distribution under the null. Generally, to construct the valid bootstrap null distribution, we shall impose the restriction that $\theta = 1$ into our re-sampling procedures.

4.2 Algorithm

1. Given a sample $\{y_t\}_{t=1}^{T}$, generated from (**??**), and (**??**): where $m = 0$ or 1, ϵ_t is IID (0, σ_ϵ^2), with unknown distribution F. Let δ be parameters to be estimated, When $m = 0$ and $m = 1$, $\delta = (\beta_1, \alpha, F)$ and $\delta = (\alpha, F)$, respectively. δ is estimated using the maximum likelihood principle:

   (a) i. If $m = 0$,

   $$\Delta y_t = \alpha \Delta y_{t-1} + \eta_t - \theta \eta_{t-1}, \quad (12)$$

   ii. If $m = 1$,

   $$\Delta y_t = \beta_1 + \alpha \Delta y_{t-1} + \eta_t - \theta \eta_{t-1}, \quad (13)$$

   which gives the estimates for $\hat{\beta}_1$, $\hat{\alpha}$ and residuals $\hat{\eta}_t$.

   (b) Center the residual $\hat{\eta}_t$ by:

   $$\bar{\eta}_t \equiv \hat{\eta}_t - \frac{1}{T-1} \sum_{t=2}^{T} \hat{\eta}_t \quad (14)$$

   (c) Re-sample without replacement from the empirical distribution unction $\{\bar{\eta}_t\}$, denoted by $\bar{F}_T$, the estimate o F. The re-draw is of size T, and denoted by $\{\eta_t^*\}$.

2. Set the initials that $y_1^* = y_1$, $y_2^* = y_2$, and generate the bootstrap samples $\{y_t^*\}$, based on:

$$y_t^* = y_{t-1}^* + \hat{\alpha} \Delta y_{t-1}^* + \eta_t^* - \eta_{t-1}^* \quad (m = 0); \quad (15)$$

$$y_t^* = y_{t-1}^* + \hat{\beta}_1 + \hat{\alpha} \Delta y_{t-1}^* + \eta_t^* - \eta_{t-1}^* \quad (m = 1) \quad (16)$$

3. Calculate $KPSS_\mu(k)$ and $KPSS_\tau(k)$ using $\{y_t^*\}_{t=1}^{T}$, denoted by $KPSS_\mu^*(k)$ and $KPSS_\tau^*(k)$, respectively.

4. Repeat step 1(c) to step 3 NB time. This constructs the bootstrap null distribution for $KPSS_\mu^*(k)$ and $KPSS_\tau^*(k)$.

5. Compute the bootstrap critical values, based on the bootstrap null distribution.

5 Self Evaluation of the Project

The project started with an idea that was to propose a bootstrap scheme to improve the size distortions in testing for possible breaks in cointegrating relation. The idea proved to be too general and too ambitious. Tests for structural changes in cointegrating vector in fact build on the same structure of tests for statioanrity null. The size problem that stationartiy tests have are also facing tests for structural breaks in cointegration. Without working through the size problem with stationarity tests, it is infeasible to work out the case with tests for structural changes in coiniegration. We turn our attention to the case with stationarity tests, and found indeed that there have been subtle difficulties that have not been foreseen in the beginning of making the NSC proposal. The current report summarizes some major results from working on stationarity tests. It shows some promises for the results to be published in international journals, if well written. In particular, the bootstrap stationarity testing procedure has not seen in the literature, and can be considered some new in the literature.

References

1. Andrew, D.W.K. (1991) "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation." *Econometrica* 59, 817-858.

2. Caner, M. and L., Kilian (2001) "Size Distortions of Tests of the Null Hypothesis of Stationarity: Evidence and Implications for the PPP debate." *Journal of International Money and Finance* 20, 639-657.

3. Elliott, G., T. J. Rothenberg and J. H. Stock (1996) "Efficient Tests for An Autoregressive Unit Root." *Econometrica* 64, 813-836.

4. Engel, C. (2000) "Long-Run PPP May Not Hold After All." *Journal of International Economics* 51, 243-273.

5. Engel, C. and C.J. Kim (1999) "The Long-Run U.S.-U.K. Real Exchange Rate." *Journal of Money, Credit, and Banking* 31, 335-356.

6. Horowitz, L. (2000) "The Bootstrap." manuscript, University of Iowa.

7. Kuo, B.-S. and A. Mikkola (1999) "Re-examining Long-run Purchasing Power Parity." *Journal of International Money and Finance* 18, 251-266.

8. Kwiatkowski, D., P.C.B. Phillips, P. Schmidt, and Y. Shin (1992) "Testing the Null 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.

9. Lee, J. (1996) "On the Power of Stationarity Tests Using Optimal Bandwith Estimates." *Economics Letters* 51, 131-137.

10. Nabeya, S. and K. Tanaka (1988) "Asymptotic Theory of a Test for the Constancy of Regression Coefficients Against the Random Walk Alternative." *Annals of Statistics* 16, 218-235.

11. Nankervis, J. & N. Savin (1996) "The Level and Power of the Bootstrap t Test in the AR(1) Model With Trend." *Journal of Business & Economic Statistics* 14, 161-168.

12. Newey, W.K. and K.D. West (1987) "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55, 703-708.

13. Ng, S. and P. Perron (2001) “Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power.” *Econometrica* 69, 1519-1554.

14. Park, J.Y. (2002) “An Invariance Principle for Sieve Bootstrap in Time Series.” *Econometric Theory* 18, 469-490.

15. Perron, P. and S. Ng (1996) “Useful Modifications to Some Unit Root Tests with Dependent Errors and Their Local Asymptotic Properties.” *Review of Economic Studies* 63, 435-465.

16. Potscher, B.M. (1991) “Non-invertibility and Pseudo-Maximum Likelihood Estimation of Mis-specified ARMA Model.” *Econometric Theory* 7, 435-449.

17. Psaradakis, Z. (2001) “Bootstrap Tests for an Autoregressive Unit Root in the Presence of Weakly Dependent Errors.” *Journal of Time Series Analysis* 22, 577-594.

18. Schwert, G.W. (1989) “Tests for Unit Roots: A Monte Carlo Investigation.” *Journal of Business and Economic Statistics* 7, 147-160.