A Gaussian test for unit roots with an application to great ratios

Abeysinghe, Tilak; Rajaguru, Gulasekaran

Published in:
52nd Annual Conference of the NZ Association of Economists Proceedings

Published: 01/01/2011

Document Version:
Publisher's PDF, also known as Version of record

52nd NZAE Conference, Wellington

A Gaussian Test for Unit Roots with an Application to Great Ratios

Tilak Abeysinghe* and Gulasekaran Rajaguru**

*Department of Economics
National University of Singapore
Email: tilakabey@nus.edu.sg

**School of Business
Bond University
Australia
Email: rgulasek@bond.edu.au

Abstract

Non-standard distributions are a common feature of many tests for unit-roots and cointegration that are currently available. The main problem with non-standard distributions is that when the true data generating process is unknown, which is the case in general, it is not easy to engage in a specification search because the distribution changes as the specification changes, especially with respect to deterministic components. We use a mixed-frequency regression technique to develop a test for cointegration under the null of stationarity of the deviations from a long-run relationship. What is noteworthy about this MA unit root test, based on a variance-difference, is that, instead of having to deal with non-standard distributions, it takes testing back to the normal distribution and offers a way to increase power without having to increase the sample size substantially. Monte Carlo simulations show minimal size distortions even when the AR root is close to unity and that the test offers substantial gains in power against near-null alternatives in moderate size samples. Although the null of stationarity is the research line to be pursued, we also consider an extension of the procedure to cover the AR unit root case that provides a Gaussian test with more power. An empirical exercise illustrates the relative usefulness of the test further.

Key words: Null of stationarity, MA unit root, mixed-frequency regression, variance difference, normal distribution, power.

1. Introduction

Unit root tests, though used extensively in applied work, are beset with problems of non-standard distributions, size distortions and extremely low power.[1] The biggest problem with non-standard distributions is that when the true data generating process is unknown, which is the case in general, it is not easy to engage in a specification search because the distribution changes as the specification changes, especially with respect to deterministic components. As Cocharan (1991, p. 202) expressed: "To a humble *macroeconomist* it would seem that an edifice of asymptotic distribution theory that depends crucially on unknown quantities must be pretty useless in practice." Some reprieve to this has been offered by Phillips (1998, 2002) who argued that there is no point of considering the trend-stationary alternative because the limiting forms of unit root processes can be expressed entirely in terms of deterministic trend functions.

In this exercise we are re-visiting the problem with the objective of presenting a unit root and cointegration test based on the null of stationarity that put the distribution back to Normal and offers substantial improvement in size and power properties. The importance of tests based on the null of stationarity need not be overemphasized. Unit roots in individual series are not that much of interest to economists. What is of interest is whether the regression provides stable parameters with stationary residuals regardless of the nature of the non-stationarity of the individual series. For example, two variables which are causally related may have structural breaks in them and AR unit root tests may take them to be I(1) processes. In a regression relationship, however, the structural break

[1] See Maddala and Kim (1998) for an extensive survey of the unit root literature.

may disappear and the regression may render stationary residuals.[2] Therefore, forming a null of stationarity will allow us to test it against different alternatives such as AR unit roots, fractional integration, structural breaks and policy interventions. The relevant alternative has to depend on the particular empirical analysis carried out. In this exercise we consider only the AR unit root alternative and defer the evaluation of other alternatives to future work. It should be noted, however, that, as the literature on structural breaks highlights, an AR unit root could be a manifestation of the misspecification of the basic regression relationship; therefore the AR unit root alternative encompasses many other forms of non-stationarities.

The basic concept embodied in our test procedure emanated from a mixed-frequency regression presented in Abeysinghe (1998, 2000) and temporal aggregation and dynamic relationships studied in Rajaguru and Abeysinghe (2002) and Rajaguru (2004). The test procedure involves a simple data transformation to obtain a mixed frequency regression and focusing on the difference in error variances of the original model and the transformed model. This method can be exploited to develop even better tests with standard distributions. As an extension, we consider the non-stationary null (AR unit root) as well.

2. Power of Existing Unit Root Tests

Table 1 provides a non-exhaustive summary of power of unit root tests near the null at sample size 100 (or 200 in a few cases). Panel (a) in the table is for the non-stationary null (AR unit root) and panel (b) for the stationary null (MA unit root or its variants).

[2] There are also cases where economic theory leads to using variables like investment/GDP ratio or the average tax rate in regressions. The meaning of a unit root in these variables is unclear.

Panel (a) also includes a representative citation of power under structural breaks. The literature on unit roots under structural breaks has also grown rapidly and we do not digress into this literature. The reference model given in the table involves an over-simplification for some simulation exercises. A general specification of the stationary null is given in models (1) and (2) of the paper.

The summary in Table 1 highlights the low power of unit root tests in general though some test procedures produce reasonably large power at sample size 100. As stated earlier, all these tests have to deal with non-standard distributions and increasing power needs increasing the sample size. These are the problems that we try to overcome in this exercise.

==============

Insert Table 1 here

==============

3. Methodology

Consider the following model that Leybourne and McCabe (1994) extended from Harvey (1989) and Kwiatkowski et al. (1992) to test the null of stationarity against an alternative of difference stationarity:

$$\begin{aligned} \phi(L)y_t &= \alpha_t + \beta t + \varepsilon_t \\ \alpha_t &= \alpha_{t-1} + \eta_t, \ \alpha_0 = \alpha \end{aligned} \quad (1)$$

where $\varepsilon_t \sim iid(0,\sigma_\varepsilon^2)$, $\eta_t \sim iid(0,\sigma_\eta^2)$, both of which are independent of each other, and $\phi(L) = 1 - \phi_1 L - ... - \phi_p L^p$ with roots outside the unit circle. This has the following ARIMA(p,1,1) representation:

$$\Delta y_t = \beta + \phi_1 \Delta y_{t-1} + ... + \phi_p \Delta y_{t-p} + u_t - \theta u_{t-1} \quad (2)$$

where $u_t \sim iid(0,\sigma^2)$ with $\sigma^2 = \sigma_\varepsilon^2 / \theta$, $\theta = (\lambda - (\lambda^2 + 4\lambda)^{1/2} + 2)/2$ and $\lambda = \sigma_\eta^2 / \sigma_\varepsilon^2$ is the signal-to-noise ratio. The so-called hyper-parameter σ_η^2 is a measure of the size of the random walk in (1). If $\sigma_\eta^2 = \lambda = 0$, $\theta = 1$ and model (2) collapses to a stationary AR(p) process. Alternatively, Δy_t in (2) has a non-invertible ARMA(p,1) representation. To test the null of stationarity a number of researchers formulated tests based on $H_0 : \sigma_\eta^2 = 0$ vs $H_1 : \sigma_\eta^2 > 0$. These are in effect tests of the MA unit root and the distributions involved are in general non-standard.

As λ increases, θ approaches zero and we get a standard unit root autoregression:

$$y_t = \rho y_{t-1} + \beta + \phi_1 \Delta y_{t-1} + ... + \phi_p \Delta y_{t-p} + u_t \quad (3)$$

with $\rho = 1$.

In our paper the ARIMA model in (2) forms the basis for our main test and for this reason we denoted σ_u^2 by σ^2. We then extend the test procedure for H_0: $\rho = 1$ in (3). This provides an AR unit root test with more power.

3.1 Null of Stationarity (MA Unit Root)

As stated earlier our test is based on a mixed frequency regression procedure (Abeysinghe, 1998, 2000) that helps in increasing the power of the test at a given sample size. To illustrate the idea, (2) can be written as

$$u_t = \theta u_{t-1} - \beta + \phi(L) \Delta y_t . \quad (4)$$

If u_t is assumed to be observed at intervals $t = m, 2m, ..., T$, where $m \geq 2$ is a positive integer, and Δy_t is observed at intervals $t = 1, 2, .., T$, the basic idea of the mixed frequency regression is to transform u_{t-1} in (4) to u_{t-m}. This transformation is easily obtained by multiplying (4) through by the polynomial $\theta(L) = 1 + \theta L + ... + \theta^{m-1} L^{m-1}$. The transformed model can be written as

$$\theta(L)\phi(L)\Delta y_t = \theta(1)\beta + V_t \qquad (5)$$

where $V_t = \theta(L)(1-\theta L)u_t = u_t - \theta^m u_{t-m}$.

Now note that under the null $H_0 : \theta = 1,\ Var(V_t) = \sigma_m^2 = 2\sigma^2$ and under the alternative $H_1 : |\theta| < 1,\ Var(V_t) = (1+\theta^{2m})\sigma^2 < 2\sigma^2$. Therefore, $\sigma_m^2 - 2\sigma^2$ forms the basis of our test. By transforming the test of θ into a test of $Var(V_t)$ we can arbitrarily increase the distance between the null and the alternative simply by increasing *m* whereby a substantial gain in power is made possible. For example, a test of $\theta = 1$ when $\theta = 0.9$ translates into comparing $2\sigma^2$ against $Var(V_t) = 1.43\sigma^2$ for *m*=4 and $Var(V_t) = 1.08\sigma^2$ for *m*=12. This transformation allows us to formulate a number of test statistics that follow standard distributions and thus alleviates a serious handicap of current unit root tests.

Given that we can obtain consistent estimates of the parameters in (2), we can compute $\hat{\sigma}^2$ and $\hat{\sigma}_m^2$ (see below) and then form the test statistic $\sqrt{T}\,(\hat{\sigma}_m^2 - 2\hat{\sigma}^2)$ to test $\theta = 1$ against $|\theta| < 1$. To establish the distribution of the test statistic define the $T \times (1+p)$ design matrix ***X*** with the *t*th row given by $(1, \Delta y_{t-1}, ..., \Delta y_{t-p})$ and the $(1+p)\times 1$ vector ***c*** given in the text before (A8). Using the subscript T to indicate the dependence on the sample size the following theorem establishes the asymptotic distribution of the test statistic.

Theorem 1

Given that $u_t \sim iid(0,\sigma^2)$ and assuming $E(u_t^4)=\mu_4<\infty$, under the null hypothesis of $\theta=1$, $\sqrt{T}\,(\hat{\sigma}^2_{m,T}-2\hat{\sigma}^2_T)\xrightarrow{d} N(0,4\sigma^4)$. In small samples

$$Var[\sqrt{T}\,(\hat{\sigma}^2_{m,T}-2\hat{\sigma}^2_T)]=4[\sigma^4+\mu_4(T/(T-m)-1)-2(\mu_4-\sigma^4)(mT/(T-m)^2) \\ -\sigma^2\boldsymbol{c}'(\boldsymbol{X}'\boldsymbol{X}/T)^{-1}\boldsymbol{c}].$$

Proof: see Appendix.

Our Monte Carlo simulation exercise shows that when $m>p$, $\sigma^2\boldsymbol{c}'(\boldsymbol{X}'\boldsymbol{X}/T)^{-1}\boldsymbol{c}$ contributes very little to the variance and can be ignored. The test procedure in practice, therefore, is the following. Assuming p+1 pre-sample values $y_{-p},...,y_0$ are available, estimate the ARMA(p,1)[3] for Δy_t in (2) by ML and obtain $\hat{\theta}$ and $\hat{\sigma}^2=\sum_{t=1}^{T}\hat{u}_t^2/(T-p-2)$ (these are provided by standard computer software procedures). Then obtain $\hat{V}_t=\hat{u}_t-\hat{\theta}^m\hat{u}_{t-m}$ and $\hat{\sigma}^2_m=\sum_{t=m+1}^{T}(\hat{V}_t-\bar{\hat{V}})^2/(T_a-1)$, where $T_a=T-m$, and compute the z score. If u_t is assumed to be Normal then $z=\sqrt{T}(\hat{\sigma}^2_m-2\hat{\sigma}^2)/[2\hat{\sigma}^2(1+3(T/T_a-1)-mT/T_a^2)^{1/2}]$ and reject the null

[3] Model selection criteria and the usual diagnostics may be used for determining the structure of the model (see the empirical exercise in Section 5).

hypothesis $\theta = 1$ if $z \le c$ where *c* is a left-hand critical value from the standard normal distribution. We may term this as z(MA) test.

In estimating θ there are two problems that we have to guard against. One is the well known pile-up problem of the ML estimator at the invertibility boundary (see Breidt et al., 2006, for references). The pile-up problem is an issue that is being addressed by a number of researchers. In particular Davis and Dunmuir (1996) have explored the possibility of using a Laplace likelihood with a local maximizer to estimate an MA(1) model with a unit root or a near unit root. It is very likely that an estimator of θ that will overcome the pile-up problem will emerge in due course. From a practical point of view, the pile-up problem of the Gaussian likelihood may not be a serious problem. Although over-differenced stationary series produce $\theta = 1$, AR unit-root series are likely to produce a θ well away from unity. Many empirical estimates of θ from non-stationary series hardly exceed 0.9 and do not exhibit the presence of the pile-up problem. As we shall see, our test offers sufficient power against the alternative of θ=0.9 in moderate-sized samples.

The other difficulty is the near common factor problem. Although the ML estimator of θ under the null is T-consistent (see Davis and Dunmuir, 1996, and reference therein), an AR factor with a root close to unity may render a highly unreliable estimate of θ in certain samples. The near common factor problem can easily be spotted by fitting an AR(p) model to y_t and ARMA(p,1) to Δy_t (see the application in Section 5). If y_t is stationary with an AR root near unity and if it is not well estimated in the ARMA model then it is important to re-estimate the model using different starting values for θ, including $\theta=1$.[4]

3.2 Monte Carlo results

In this section we present the results of a limited number of Monte Carlo experiments to highlight the size and power properties of the test under near unit root alternatives. The basic generating process we consider is the following ARIMA(1,1,1) model:

$$(1-\phi_1 L)\Delta z_t = \beta + (1-\theta L)u_t\,. \tag{6}$$

We consider two cases. In the first set up, z_t represents a single data series that is being tested for unit roots or cointegration with a known coefficient vector. In the second set up z_t represents OLS residuals from a static regression that is being tested for cointegration. The cointegrating model is given by

$$\begin{aligned} y_t &= \delta_0 + \delta_1 x_t + z_t \\ x_t &= x_{t-1} + \varepsilon_t. \end{aligned} \tag{7}$$

In this setting, both u_t and ε_t are generated from independent $N(0,1)$ distributions. The size of the test is obtained when $\theta=1$. For this we set $\phi_1 = 0.5,\ 0.9,\ 0.95$. For power, we use $\theta = 0.8,\ 0.9$ with $\phi_1 = 0.5$. In the first setting $\beta = 1$ and in the second setting $\beta = 0$ and $\delta_0 = \delta_1 = 1$. To obtain a preliminary assessment we conducted the simulation experiment for *T*=300 and *m*=2,4,6,8,10,12 and observed that as *m* increases the size also tends to increase slightly. For example, at the 5% level a representative sample of

---

[4] It is instructive to use a dedicated ARIMA software procedure for estimation. We used SAS PROC ARIMA in our exercise by removing the default boundary constraint by invoking "nostable" option. One could devise alternative estimators that avoid the need for estimating θ. We tried an IV estimator, but it did not improve power much.

rejection frequencies for *m* given above is (0.032,0.058,0.072,0.083,0.075,0.083). Since the test relies on the consistency of $\hat{\theta}$, small-sample bias of the estimator tends to distort the distribution of the test statistic as *m* increases. Based on both size and power an optimal choice of *m* seems to be 4 for moderately sized samples. Table 2 reports detailed results for *m*=4. What is important to note in this table is that testing regression residuals for cointegration does not lead to much distortion in size or a reduction in power.

Table 2: Size and power of MA unit root test for *m*=4
(z(MA) test, 2000 replications)

T	Single series (known cointegrating vector)			Regression residuals		
	1%	5%	10%	1%	5%	10%
Size						
ϕ_1=0.5, θ=1						
100	0.021	0.045	0.075	0.046	0.084	0.129
200	0.015	0.046	0.079	0.022	0.064	0.104
300	0.022	0.064	0.103	0.023	0.056	0.090
500	0.014	0.055	0.098	0.012	0.048	0.095
ϕ_1=0.9, θ=1						
200	0.023	0.070	0.107	0.040	0.087	0.132
300	0.027	0.077	0.128	0.028	0.078	0.122
500	0.008	0.043	0.096	0.019	0.060	0.105
ϕ_1=0.95, θ=1						
500	0.014	0.046	0.081	0.025	0.072	0.121
Power						
ϕ_1=0.5, θ=0.8						
100	0.572	0.609	0.637	0.507	0.547	0.572
200	0.874	0.895	0.905	0.827	0.848	0.862
300	0.957	0.967	0.974	0.947	0.957	0.960
500	0.995	0.996	0.997	0.993	0.994	0.994
ϕ_1=0.5, θ=0.9						
100	0.315	0.357	0.391	0.242	0.293	0.326
200	0.606	0.674	0.705	0.555	0.623	0.658
300	0.809	0.848	0.875	0.794	0.834	0.852
500	0.955	0.964	0.968	0.950	0.959	0.963

4. Extension: Null of Non-stationarity (AR unit root)

We can extend the above procedure to transform the ADF test to a Gaussian test based on the variance difference. Now the basic regression model is the one in (3). The mixed frequency regression procedure in this case involves multiplying (3) through by the polynomial $\rho(L) = 1 + \rho L + ... + \rho^{m-1} L^{m-1}$ that yields

$$\rho(L)(1-\rho L) y_t = \rho(1)\beta + \phi_1 \rho(L) \Delta y_{t-1} + ... + \phi_p \rho(L) \Delta y_{t-p} + \rho(L) u_t . \qquad (8)$$

This provides

$$y_t = \rho^m y_{t-m} + \beta \sum_{i=0}^{m-1} \rho^i + \phi_1 \sum_{i=0}^{m-1} \rho^i \Delta y_{t-1-i} + ... + \phi_{p-1} \sum_{i=0}^{m-1} \rho^i \Delta y_{t-p+1-i} + V_t \qquad (9)$$

where $V_t = \sum_{i=0}^{m-1} \rho^i u_{t-i}$. Now under the null $H_0 : \rho = 1$, $Var(V_t) = \sigma_m^2 = m\sigma^2$ and under the alternative $H_1 : |\rho| < 1$, $Var(V_t) = \sigma^2 \sum_{i=0}^{m-1} \rho^{2i} < m\sigma^2$. In this case, $\sigma_m^2 - m\sigma^2$ forms the basis of our test. Note that for notational consistency we have re-defined V_t and σ_m^2. As in the MA unit root case, by transforming the test of ρ into a test of $Var(V_t)$ we can arbitrarily increase the distance between the null and the alternative simply by increasing *m* whereby a substantial gain in power is made possible. For example, a test of $\rho = 1$ when $\rho = 0.9$ translates into comparing $m\sigma^2$ against $Var(V_t) = 3.44\sigma^2$ for *m*=4 and $Var(V_t) = 7.18\sigma^2$ for *m*=12.

To establish the distribution of $\sqrt{T}\,(\hat{\sigma}_m^2 - m\hat{\sigma}^2)$, define the $T \times (1+p)$ design matrix $\boldsymbol{X}$ with tth row given by $(1, \Delta y_{t-1}, ..., \Delta y_{t-p})$ and $(1+p)\times 1$ vector $\boldsymbol{c}$ given in A21. Using the subscript T to indicate the dependence on the sample size, the following theorem establishes the asymptotic distribution of the test statistic $\sqrt{T}\,(\hat{\sigma}_m^2 - m\hat{\sigma}^2)$.

Theorem 2

Given that $u_t \sim iid(0, \sigma^2)$ and assuming $E(u_t^4) = \mu_4 < \infty$, under the null hypothesis of $\rho = 1$ $\sqrt{T}(\hat{\sigma}_{m,T}^2 - m\hat{\sigma}^2) \xrightarrow{d}$

$N(0, 2m(2m-1)(m-1)\sigma^4/3) - 4\sigma^2 \boldsymbol{c}'(\boldsymbol{X}'\boldsymbol{X}/T)^{-1}\boldsymbol{c})$.

Proof: See Appendix.

To avoid the possibility of a negative estimate for the variance in the above theorem we recommend setting $m \gg p$. When p=0, $\boldsymbol{c}$ is a zero vector. Incidentally when p=0 the variance of the above test statistic specializes to that of the variance ratio test obtained by Lo and MacKinlay (1988) in a different setting under the *iid Gaussian* assumption. The major advantage of our test over the variance ratio test is that our test entails a mechanism to increase power at a given sample size.

The test procedure in practice is the following. Assuming p+1 pre-sample values $y_{-p}, ..., y_0$ are available, obtain the OLS estimate $\hat{\rho}$ and the residuals $\hat{u}_t$ from the regression in (3) and compute $\hat{\sigma}^2 = \sum_{t=1}^{T} \hat{u}_t^2 / (T - (2+p))$, $\hat{V}_t = \sum_{i=0}^{m-1} \hat{\rho}^i \hat{u}_{t-i}$ and

$\hat{\sigma}_m^2 = \sum_{t=m}^{T} (\hat{V}_t - \bar{\hat{V}})^2 / n$, where $n = (T - m + 1)(1 - 2m/T)$ (see Appendix). Then compute $z = \sqrt{T}(\hat{\sigma}_m^2 - m\hat{\sigma}^2)/\sqrt{\text{var}}$, where var is the variance given in Theorem 2 with σ replaced with $\hat{\sigma}$, and reject the null hypothesis $\rho = 1$ if $z \le c$ where c is a left-hand critical value from the standard normal distribution. We may term this as z(AR) test.

Monte Carlo experiments as in Section 3.2 show that in small samples such as T=100 the test produces desirable size properties when p=0 in (3) and some size distortions occur when p≥1. Even when p=0 size remains close to the nominal level as m increases only if the test is treated as a two-tailed test; as m increases the left-tail probability declines and the right-tail probability increases while the sum remains the same. This is also what Lo and MacKinlay (1989) observed in their variance ratio test and they used a two-tailed test. This problem, however, disappears as the sample size increases. One-tailed test can be used safely in samples of size about 100 by setting $m \le 8$. The test entails substantial gains in power. For example, when ρ=0.9 and T=100, the power of the 5% one-tailed test increases from 0.24 to 0.45 when m is increased from 2 to 8. At this stage we can recommend the test for p=0 setting only.[5]

5. Some empirical results

As empirical illustrations we present two tables of results, the first is a representative set of variables from Abeysinghe and Choy (2007) where they present a 62-equation macroeconometric model (ESU01 model) for the Singapore economy, the second is a test

[5] Since the test requires further work for higher order AR processes we do not report the Monte Carlo results to conserve space.

of stationarity of the average propensity to consume (APC) in OECD countries and finally analyze the unit root properties for *ln*(*I/Y*).

Abeysinghe and Choy (2007) estimated all the key behavioral equations in their model individually in the form of error correction models by crafting out the underlying long-run (cointegrating) relationships carefully paying attention to specific features of the Singapore economy, economic theory, and parameter stability. Table 3 presents test results for two groups of cointegrating relationships: (i) cointegrating regression residuals[6] and (ii) relations with known coefficients. In the latter group, the oil price equations were designed to check the extent of exchange rate pass-through.[7] Relative unit business cost (RUBC) and the real exchange rate (RER) are both measures of competitiveness. Although the RER presented in the table is not a variable in the ESU01 model we use it here for further illustration of the performance of the test.

In Table 3, all series except for RER clearly pass as AR(1) processes and it is worth noting that the estimates of ρ from AR and ARMA(1,1) models are very close. Therefore, first estimating an AR(p) model provides a good check against the ARMA(p,1) estimation for the MA unit root test. It is also useful to note that when over-differencing is not involved as in the RER case (also those in Table 4 below) the MA root is likely to be a distance away from unity in many practical cases and as a result our test carries a lot of power against such alternatives.

---

6 Readers interested in the regression equations are referred to Abeysinghe and Choy (2007).

7 As the third largest oil refining center and trading hub in the world Singapore may have some price setting power on its oil market in which case the stationarity of the long-run relationship with unity restriction has to be rejected. Note that short-run pass through is well below unity.

The test results in Table 3 show that if we were to use the ADF test to test for cointegration only three equations (consumption, exports and oil export price) qualify as cointegrating relationships (the null of AR unit root is rejected). Our z(MA) test, on the other hand, does not reject the null of stationarity (and cointegration) in all the cases except the last one. The RER series with $\hat{\rho}$ =0.98 clearly comes out as a non-stationary process. Unlike the ADF test, our z(AR) test concurs with the outcome of the z(MA) test with one exception, the CPI equation. As stated earlier z(AR) test needs further refinements. Since Abeysinghe and Choy (2007) have already studied these cointegrating relationships in detail and that the z(MA) test concurs with these findings is a strong case in favor of the new test.

Table 3: Cointegration test for selected equations
from the ESU01 model of the Singapore economy
(Abeysinghe and Choy, 2007)

					Variance Difference	
Equation in the model	T	$\hat{\rho}$	ARMA(1,1)	ADF	z(MA) m=4	z(AR) m=8
(i) Regression Residuals						
Consumption	104	0.67	0.70, 0.99	-4.48*	-0.77	-2.65*
Exports (non oil domestic)	96	0.54	0.56, 0.99	-5.27*	0.63	-2.71*
Employment	96	0.86	0.88, 0.99	-2.41	0.51	-1.73*
Wages	96	0.89	0.87, 0.99	-2.94	0.49	-1.93*
CPI	96	0.93	0.95, 0.99	-2.01	0.05	-0.86
(ii) Known coefficients (log form)						
Oil import price in S$	104	0.89	0.85, 0.99	-2.43	-1.49	-2.92*
Oil export price in S$	104	0.76	0.79, 0.99	-3.68*	0.42	-2.99*
RUBC	96	0.91	0.93, 0.99	-2.17	0.25	-1.72*
RER	336	0.98	0.00, -0.25	-2.39	-9.03*	-1.20

RUBC=relative unit business cost. RER=real exchange rate (S$/US$, CPI based). Oil price relationships are: oil price in Singapore dollars equals oil price in US$ times the Sin/US exchange rate. First eight series are quarterly from 1978Q1 or 1980Q1 to 2003Q4. RER is monthly over 1975-2003. The null for z(MA) is stationarity (MA unit root) and that for ADF and z(AR) is non-stationarity (AR unit root). * significant at the 5% level (left-tail test).

As a further illustration of the test, Table 4 presents the test results on APC for 21 OECD countries.[8] Because of the non-availability of sufficiently long data series on non-durable consumption and disposable income we measure APC by the ratio of total consumption expenditure to GDP. Although the APC is expected to be stationary for developed economies, some countries show local trends in their APCs over the sample period. This is reflected in large values of $\hat{\rho}$ (the sum of AR coefficients) in Table 4. This is where many tests may misconstrue APC to be an I(1) process.

As in Table 3 we can notice the close correspondence between AR(p) coefficients and ARMA(p,1) coefficients in identifiable stationary cases. It is also worth noticing that in stationary cases $\hat{\theta}$ turns out to be almost unity. This means that the size distortion we noticed in the Monte Carlo experiment resulting from under estimation of θ may not be a serious problem in practice.

Again the ADF test turns out to be the least powerful against near unit root alternatives, it renders the I(1) verdict on 18 of the 21 APC series. The Johansen test fairs reasonably better, it recognizes eight cases as cointegrating relationships. Our z(MA) test on the other hand, takes 16 of the APC series to be stationary. It rejected stationarity only when $\hat{\rho} \geq 0.97$ and when the local trend dominated the series; see the cases of Canada and Korea for a comparison, both with $\hat{\rho} = 0.97$, one is assessed to be I(0), the other I(1). Like many fast growing developing economies Korea experienced a falling APC till the mid

[8] Data for this exercise are from the IFS database except for France. IFS data for France show some irregularities, therefore, France data were taken from the OECD database which covers a shorter time span than the IFS database.

1980s before stabilizing to fluctuate around a constant mean. Rejecting the null of stationarity of APC is, therefore, an indication of the interplay of other variables that need to be considered instead of taking APC to be an I(1) process.

Table 4: Cointegration test on APC

Country	Sample period (quarterly)	T	AR Lags	AR Coefficients	$\hat{\rho}$	ARMA(p,1)	ADF	Johansen VAR(4)	z(MA) m=4
Australia	1960-2007	192	1	0.92	0.92	0.94, 0.99	-2.71	yes	0.39
Austria	1965-2007	172	1,2,3	0.55,0.18,0.18	0.91	0.56, 0.19, 0.20, 0.99	-2.33	no	0.34
Belgium	1980-2007	111	1	0.98	0.98	0.00, 0.12	-0.77	no	-5.37*
Canada	1957-2007	204	1	0.97	0.97	0.97, 0.99	-1.97	no	-0.30
Denmark	1978-2007	124	1,4	0.75, 0.21	0.96	0.75, 0.17, 0.99	-1.71	yes	-0.57
Finland	1970-2007	152	1,4	0.71, 0.21	0.92	0.72, 0.19, 0.99	-2.21	no	-1.41
France	1978-2007	120	1	0.94	0.94	0.97, 0.99	-2.1	yes	0.48
Germany	1961-2007	188	1,3	0.71, 0.23	0.94	0.72, 0.23, 0.99	-1.99	yes	-1.13
Italy	1970-2007	151	1,4	0.70, 0.12	0.82	0.66, 0.99	-2.98*	yes	-0.5
Japan	1965-2007	172	1	0.94	0.94	0.95, 0.99	-2.45	no	-1.29
Korea, South	1965-2007	172	1	0.97	0.97	0.00, 0.20	-2.45	no	-6.51*
Mexico	1981-2007	108	1	0.88	0.88	0.88, 0.99	-2.62	no	-0.14
Netherlands	1977-2007	124	1,2	0.51, 0.46	0.97	0.35, 0.25	-0.78	no	-5.75*
New Zealand	1987-2007	82	1	0.72	0.72	0.75, 0.99	-3.69*	yes	0.58
Norway	1961-2007	188	1,2	0.75, 0.23	0.98	0.00, 0.25	-0.83	no	-6.52*
Spain	1970-2007	152	1,4	0.79, 0.20	0.99	0.00, 0.24	-0.06	no	-6.66*
Sweden	1980-2007	112	1,2,4	0.66, 0.39, -0.17	0.88	0.61, 0.41, -0.17, 0.99	-2.21	no	-0.81
Switzerland	1970-2007	152	1,2,3	0.60, 0.51, -0.18	0.94	0.59, 0.53, -0.16, 0.99	-1.81	no	-1.11
Turkey	1987-2007	83	1	0.62	0.62	0.57, 0.99	-4.23*	yes	0.36
UK	1957-2007	204	1,3	0.73, 0.24	0.97	0.73, 0.25, 0.99	-1.55	yes	1.21
US	1957-2007	204	1,2	0.83, 0.17	1.00	0.00, 0.17	-0.18	no	-7.29*

Note that some data series end in Q2 or Q3 in 2007. Tests are based on log(APC) = log(C/Y), where C is total consumption expenditure and Y is GDP, both in nominal terms and seasonally adjusted. * Significant at the 5% level (left-tail test). For the Johansen test "yes" means acceptance of cointegration between log(C) and log(Y) with the cointegrating vector (1, -1).

The results reported in Table 5 for 25 OECD countries further assures that unit root hypothesis is rejected for most of the countries based on the proposed test than the other traditional alternatives.

Table 5. Unit Root Test on ln(Investment/GDP)

Country	Sample Period	ADF Test			PP Test	KPSS Test	Z(MA) Test		
		Lags	$\hat{\rho}$	ADF			AR Lags	$\hat{\theta}$	z(MA), m=4
Australia	1949-2010	0	0.78	-3.59**	-3.60**	0.15	0	0.98	-0.71
Austria	1948-2010	0	0.81	-4.87***	-4.82***	0.20	0	0.99	-0.53
Belgium	1953-2010	0	0.79	-2.77	-2.90	0.11	0	0.98	-0.75
Canada	1948-2010	1	0.99	-2.02	-2.05	0.12*	2	0.97	-0.48
Chile	1948-2010	1	0.39	-4.95***	5.21***	0.09	2	0.99	-0.91
Denmark	1950-2010	1	0.88	-2.11	-1.93	0.17**	1	0.96	1.30
Finland	1950-2010	2	0.84	-2.66	-2.76	0.18**	2	1.00	-0.23
France	1959-2009	5	0.84	-1.98	-2.25	0.11	5	0.95	-1.10
Germany	1960-2009	1	0.57	-3.92***	-2.56	0.04	1	0.97	1.09
Greece	1948-2009	4	0.77	-2.97**	3.56***	0.11	4	0.97	-0.77
Hungary	1970-2010	0	0.87	1.52	1.59	0.16**	0	0.07	-3.25***
Iceland	1950-2010	4	0.6	-2.97	-2.77	0.15**	4	0.62	-3.29***
Ireland	1950-2009	1	0.87	-2.51	-2.38	0.12*	2	0.95	-0.13
Italy	1951-2010	4	0.85	-2.21	-2.62	0.19**	4	0.96	-1.15
Japan	1955-2009	2	0.89	-2.37	-4.01***	0.1	0	0.96	-1.27
Korea, South	1953-2009	3	0.93	-2.25	-3.15**	0.09	4	0.99	-0.40
Luxembourg	1950-2009	2	0.63	-3.45**	-3.24**	0.33	2	0.96	0.75
Netherlands	1956-2009	2	0.77	-2.59	-2.50	0.06	3	0.96	-0.58
New Zealand	1950-2009	4	0.66	-3.68**	-3.59**	0.09	4	0.97	-1.18
Norway	1949-2010	3	0.93	-2.28	2.87	0.15**	4	0.97	-1.12
Poland	1980-2010	3	0.49	-3.28*	-3.36*	0.07	3	0.97	-0.34
Portugal	1953-2009	3	0.81	-2.71	-1.72	0.23***	3	0.95	-0.82
Spain	1954-2010	2	0.76	-3.08	-2.27	0.08	2	1.00	-0.19
Switzerland	1948-2010	2	0.85	-3.25*	-3.34*	0.10	2	0.89	-0.88
UK	1948-2009	4	0.85	-2.01	-1.94	0.25***	4	0.95	-1.47*
US	1948-2010	3	0.67	-3.97**	-3.55**	0.07	4	0.96	-0.90

*.** and *** are rejection at10%, 5% and 1% respectively

5. Conclusion

This exercise addresses three important issues. First, it highlights the importance of formulating tests based on the null of stationarity. Unfortunately the profession has not paid enough attention to this. What is of general importance is whether a regression relationship produces stationary residuals regardless of the nature of non-stationarities of the individual series. Moreover, AR unit roots in individual series is some thing hard to pin down. The apparent unit root could be a manifestation of some other forms of non-stationarity. We present an MA unit root test based on the null of stationarity. Unlike the AR unit root which is a behavioral outcome, the MA unit root is created by over-differencing and therefore easier to pin down. Although testing for an MA unit root is not new to the literature none of the existing tests have gained much popularity in applied work.

The second important aspect of the exercise is that the proposed test brings us back to Normal distribution, away from non-standard distributions, and makes specifications searches easier. The third aspect of the exercise is that the test procedure entails a mechanism to increase power without necessarily having to increase the sample size. This addresses the problem of extremely low power at near null alternatives of many unit root tests that are currently available. Despite our emphasis on the null of stationarity we also offer a test based on the null of an AR unit root that shares the above properties. Although this test requires further refinements the Monte Carlo and empirical results seem to favor the MA unit root test.

An important objection one could raise against our test is the difficulty of estimating an MA root on or near the unit circle. Some researchers are actively working on this problem and a better estimation method is likely to emerge in due course. Nevertheless, as our empirical exercise highlights, the estimation problem may not be that serious in problems encountered in practice.

Appendix

Proof of Theorem 1

Here we derive the distribution of $\sqrt{T}(\hat{\sigma}^2_{m,T} - 2\hat{\sigma}^2_T)$ under the null hypothesis $\theta = 1$. The ML estimates of the parameters are obtained by running the model in (2). Using the results below it can easily be verified that $\sqrt{T}(\hat{\sigma}^2_{m,T} - 2\hat{\sigma}^2_T) \xrightarrow{p} 0$. To derive the variance, this can be expressed as

$$\begin{aligned} E[\sqrt{T}\,(\hat{\sigma}^2_{m,T} - 2\hat{\sigma}^2_T)]^2 &= T\,E[(\hat{\sigma}^2_{m,T} - 2\sigma^2) - 2(\hat{\sigma}^2_T - \sigma^2)]^2 \\ &= T\,E[(\hat{\sigma}^2_{m,T} - 2\sigma^2)^2 + 4(\hat{\sigma}^2_T - \sigma^2)^2 - 4((\hat{\sigma}^2_{m,T} - 2\sigma^2)(\hat{\sigma}^2_T - \sigma^2)]. \end{aligned} \qquad \text{(A1)}$$

It is well established that $\hat{\sigma}^2_T = (1/T)\hat{\boldsymbol{u}}'_T\hat{\boldsymbol{u}}_T \xrightarrow{p} (1/T)\boldsymbol{u}'_T\boldsymbol{u}_T \xrightarrow{p} \sigma^2$ and $\sqrt{T}(\hat{\sigma}^2_T - \sigma^2) \xrightarrow{d} N(0,(\mu_4 - \sigma^4))$. (See, for example, Hamilton, 1994, p. 212.)

For $\hat{\sigma}^2_{m,T}$, with reference to model (2) define $\boldsymbol{\beta} = (\beta, \phi_1, ..., \phi_p)'$, the $T \times (1+p)$ matrix $\boldsymbol{X}$ with the th row given by $(1, \Delta y_{t-1}, ..., \Delta y_{t-p})$, $\boldsymbol{u} = (u_1, u_2, ..., u_T)'$, $\boldsymbol{u}_{-1} = (u_0, u_1, ..., u_{T-1})'$, $\boldsymbol{u}_{-m} = (u_{m+1}, u_{m+2}, ..., u_T)'$, and the $(T-m) \times T$ aggregation matrix $\boldsymbol{A}$:

$$A = \begin{pmatrix} \theta^{m-1} & . & \theta^2 & \theta & 1 & 0 & 0 & 0 & . & . & 0 & 0 & 0 & 0 & 0 \\ 0 & \theta^{m-1} & . & \theta^2 & \theta & 1 & 0 & 0 & . & . & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \theta^{m-1} & . & \theta^2 & \theta & 1 & 0 & . & . & 0 & 0 & 0 & 0 & 0 \\ . & . & . & . & . & . & . & . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & . & . & \theta^{m-1} & . & \theta^2 & \theta & 1 \end{pmatrix} \tag{A2}$$

Model (2) now can be written in vector-matrix notation as $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{u} - \theta \boldsymbol{u}_{-1}$. Pre-multiplying this by $\boldsymbol{A}$ and using the subscript a to indicate aggregation, we obtain $\boldsymbol{y}_a = \boldsymbol{X}_a \boldsymbol{\beta} + \boldsymbol{u} - \theta^m \boldsymbol{u}_{-m}$ which can be re-arranged to give $\boldsymbol{y}_a = \boldsymbol{X}_a \boldsymbol{\beta} + \boldsymbol{V} - (\theta^m - 1)\boldsymbol{u}_{-m}$ where $\boldsymbol{V} = \boldsymbol{u} - \boldsymbol{u}_{-m}$ under the null. Now we can obtain

$$\begin{aligned} \hat{\boldsymbol{V}} &= \boldsymbol{V} - \boldsymbol{X}_a(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) + (\hat{\theta}^m - 1)\boldsymbol{u}_{-m} \\ &= \boldsymbol{V} - \boldsymbol{X}_a^*(\hat{\boldsymbol{\beta}}^* - \boldsymbol{\beta}^*) \end{aligned} \tag{A3}$$

where $\boldsymbol{X}_a^*$ is augmented $\boldsymbol{X}_a$ with the first element of the tth row given by u_{t-m} and $(\hat{\boldsymbol{\beta}}^* - \boldsymbol{\beta}^*)$ is augmented $(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ with the first element given by $(\hat{\theta}^m - 1)$. Now defining the diagonal scaling matrix Υ of dimension $(2+p)\times(2+p)$ with the first diagonal element given by T and the rest by $T^{1/2}$ (Sims et al., 1990; Note, $\hat{\theta}$ is T-consistent) we obtain under the null:

$$
\begin{aligned}
\hat{\sigma}^2_{m,T} &= (1/T)\hat{\boldsymbol{V}}_T'\hat{\boldsymbol{V}}_T = (1/T)\boldsymbol{V}_T'\boldsymbol{V}_T - (2/T)(\hat{\boldsymbol{\beta}}^*_T - \boldsymbol{\beta}^*)'\boldsymbol{X}^{*\prime}_{aT}\boldsymbol{V}_T \\
&\qquad + (\hat{\boldsymbol{\beta}}^*_T - \boldsymbol{\beta}^*)'(\boldsymbol{X}^{*\prime}_{aT}\boldsymbol{X}^*_{aT}/T)(\hat{\boldsymbol{\beta}}^*_T - \boldsymbol{\beta}^*) \\
&= (1/T)\boldsymbol{V}_T'\boldsymbol{V}_T - 2(\hat{\boldsymbol{\beta}}^*_T - \boldsymbol{\beta}^*)'(\Upsilon_T/\sqrt{T})(\Upsilon_T^{-1}/\sqrt{T})\boldsymbol{X}^{*\prime}_{aT}\boldsymbol{V}_T \qquad \text{(A4)} \\
&\quad + (\hat{\boldsymbol{\beta}}^*_T - \boldsymbol{\beta}^*)'(\Upsilon_T/\sqrt{T})(\Upsilon_T^{-1}\boldsymbol{X}^{*\prime}_{aT}\boldsymbol{X}^*_{aT}\Upsilon_T^{-1})(\Upsilon_T/\sqrt{T})(\hat{\boldsymbol{\beta}}^*_T - \boldsymbol{\beta}^*) \\
&\xrightarrow{p} 2\sigma^2 .
\end{aligned}
$$

This result holds because $(\Upsilon_T/\sqrt{T})(\hat{\boldsymbol{\beta}}^*_T - \boldsymbol{\beta}^*) \xrightarrow{p} \boldsymbol{0}$ while the rest converge to bounded quantities.

Now we have to consider the distribution of $\sqrt{T}(\hat{\sigma}^2_{m,T} - 2\sigma^2)$. Multiplying (A4) through by $\sqrt{T}$ shows that the last term of (A4) converges in probability zero and in the second term, $T(\hat{\theta}^m - 1)(T^{-3/2}\sum X_{a1t}V_t) \xrightarrow{p} 0$ and $\sqrt{T}(\hat{\beta} - \beta)(1/T)\sum X_{a2t}V_t \xrightarrow{p} 0$. Thus we have to consider the distribution of

$$
\begin{aligned}
&\sqrt{T}(\hat{\sigma}^2_{m,T} - 2\sigma^2) \qquad \text{(A5)} \\
&\quad = \sqrt{T}(\boldsymbol{V}_T'\boldsymbol{V}_T/T_a - 2\sigma^2) - 2\sqrt{T}(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta})'(\boldsymbol{X}_{aT}'\boldsymbol{V}_T/T_a)
\end{aligned}
$$

where the presence of the constant term in the $\boldsymbol{X}_a$ matrix is inconsequential.

Now consider the variance of the first term on the RHS of (A5):

$$
\begin{aligned}
&(T/T_a^2)E[(\sum_{t=1}^{T}(V_t^2 - 2\sigma^2)]^2 \qquad \text{(A6)} \\
&\quad = (T/T_a^2)[\sum E(V_t^2 - 2\sigma^2)^2 + 2\sum E((V_t^2 - 2\sigma^2)(V_{t-k}^2 - 2\sigma^2)]
\end{aligned}
$$

where k =1,2,...

From the first term of (A6):

$$E(V_t^2 - 2\sigma^2)^2 = E(V_t^4 - 4\sigma^2 V_t^2 + 4\sigma^4) = E(V_t^4) - 4\sigma^4$$

and

$$E(V_t^4) = E[(u_t - u_{t-m})^2]^2$$
$$= E[(u_t^2 - 2u_t u_{t-m} + u_{t-m}^2)(u_t^2 - 2u_t u_{t-m} + u_{t-m}^2)]$$
$$= 2\mu_4 + 6\sigma^4.$$

Thus

$$E(V_t^2 - 2\sigma^2)^2 = 2(\mu_4 + \sigma^4).$$

From the second term of (A6):

$$E[(V_t^2 - 2\sigma^2)(V_{t-k}^2 - 2\sigma^2)] = E(V_t^2 V_{t-k}^2) - 4\sigma^4.$$

Now for $k=m$

$$E(V_t^2 V_{t-m}^2) = E[(u_t^2 - 2u_t u_{t-m} + u_{t-m}^2)(u_{t-m}^2 - 2u_{t-m} u_{t-2m} + u_{t-2m}^2)]$$
$$= \mu_4 + 3\sigma^4.$$

Thus

$$E[(V_t^2 - 2\sigma^2)(V_{t-k}^2 - 2\sigma^2)] = \mu_4 - \sigma^4, \text{ if } k = m$$
$$= 0, \text{ otherwise.}$$

Combining the two terms of (A6) we obtain:

$$(T/T_a^2)E[(\sum_{t=1}^{T}(V_t^2-2\sigma^2)]^2$$
$$\begin{aligned}&=(T/T_a^2)[T_a(2\mu_4+2\sigma^4)+(T_a-m)(2\mu_4-2\sigma^4)\\&=4\mu_4[T/(T-m)]-2(\mu_4-\sigma^4)[mT/(T-m)^2]\\&\to 4\mu_4.\end{aligned} \tag{A7}$$

Note that $(V_t^2-2\sigma^2)$ is a stationary process and therefore by the central limit theorem $\sqrt{T}(\boldsymbol{V}_T'\boldsymbol{V}_T/T_a-2\sigma^2)\xrightarrow{d}N(0,4\mu_4)$.

Now consider the second term on the RHS of (A5). To obtain its variance first note that $\sqrt{T}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})\xrightarrow{d}N(\boldsymbol{0},\sigma^2(\boldsymbol{X}'\boldsymbol{X}/T)^{-1})$ and $\boldsymbol{X}_{aT}'\boldsymbol{V}_T/T_a\xrightarrow{p}\boldsymbol{c}(\sigma^2,\phi_1,\ldots,\phi_p,m)$ where $\boldsymbol{c}$ is a $(1+p)\times 1$ vector. This vector can be derived easily by noting that the aggregated form of model (2) under the null provides, $\Delta_m y_t=\phi(L)^{-1}V_t=\psi(L)V_t$, where $\phi(L)^{-1}=\psi(L)=1+\psi_1L+\psi_2L^2+\ldots$ Note that the first term of $\boldsymbol{X}_{aT}'\boldsymbol{V}_T/T_a$ that corresponds to the constant term of the model is zero. Now consider the second term in the $\boldsymbol{X}_{aT}'\boldsymbol{V}_T/T_a$ vector:

$$\begin{aligned}(1/T_a)\sum X_{a2t}V_t&=(1/T_a)\sum\Delta_m y_{t-1}V_t=(1/T_a)\sum\psi(L)V_{t-1}V_t\\&=(1/T_a)\sum[(1+\psi_1L+\psi_2L^2+\ldots)(1-L^m)u_{t-1}](u_t-u_{t-m})\\&=-\sigma^2\psi_{m-1}.\end{aligned}$$

Proceeding in this way, we obtain for p=m: $\boldsymbol{c}'=-\sigma^2(0,\psi_{m-1},\psi_{m-2},\ldots,1)$. If p>m the $\boldsymbol{c}$ vector will have zero entries for the excess terms. Using these results the variance of the second term of (A5) can be written as

$$Var[2\sqrt{T}(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta})'(\boldsymbol{X}'_{aT}\boldsymbol{V}_T / T_a)] = 4\sigma^2 \boldsymbol{c}'(\boldsymbol{X}'\boldsymbol{X}/T)^{-1}\boldsymbol{c}. \tag{A8}$$

Now using Hausman's approach (Hausman, 1978) the variance of $\sqrt{T}(\hat{\sigma}^2_{m,T} - 2\sigma^2)$ in (A5) can be obtained as the difference of the variances given in (A7) and (A8). Thus we obtain

$$\sqrt{T}(\hat{\sigma}^2_{m,T} - 2\sigma^2) \xrightarrow{d} N(0,\ 4(\mu_4 - \sigma^2 \boldsymbol{c}'(\boldsymbol{X}'\boldsymbol{X}/T)^{-1}\boldsymbol{c})). \tag{A9}$$

Although the covariance term of (A1) can be worked out we can apply the Hausman approach again to obtain the overall variance of $\sqrt{T}(\hat{\sigma}^2_{m,T} - 2\hat{\sigma}^2)$:

$$\begin{aligned} E[\sqrt{T}(\hat{\sigma}^2_{m,T} - 2\hat{\sigma}^2_T)]^2 &= 4\mu_4 - 4\sigma^2 \boldsymbol{c}'(\boldsymbol{X}'\boldsymbol{X}/T)^{-1}\boldsymbol{c} - 4(\mu_4 - \sigma^4) \\ &= 4(\sigma^4 - \sigma^2 \boldsymbol{c}'(\boldsymbol{X}'\boldsymbol{X}/T)^{-1}\boldsymbol{c}). \end{aligned} \tag{A10}$$

This is the variance of the sum of two asymptotically normally distributed variables, hence we establish that

$$\sqrt{T}(\hat{\sigma}^2_{m,T} - 2\hat{\sigma}^2_T) \xrightarrow{d} N(0, 4(\sigma^4 - \sigma^2 \boldsymbol{c}'(\boldsymbol{X}'\boldsymbol{X}/T)^{-1}\boldsymbol{c})). \tag{A11}$$

In small samples from (A7):

$$\begin{aligned} &Var[\sqrt{T}(\hat{\sigma}^2_{m,T} - 2\hat{\sigma}^2_T)] = \\ &\quad 4[\sigma^4 + \mu_4(T/(T-m) - 1) - 2(\mu_4 - \sigma^4)(mT/(T-m)^2) \\ &\quad - \sigma^2 \boldsymbol{c}'(\boldsymbol{X}'\boldsymbol{X}/T)^{-1}\boldsymbol{c}]. \end{aligned} \tag{A12}$$

QED

Proof of Theorem 2

We need the distribution of $\sqrt{T}(\hat{\sigma}^2_{m,T} - m\hat{\sigma}^2_T)$ under the null hypothesis $\rho = 1$. The steps involved in the proof are similar to those of Theorem 1 though the quantities are different. The OLS estimates of the parameters are obtained by running the model in (3). From the results below it is easily seen that $\sqrt{T}(\hat{\sigma}^2_{m,T} - m\hat{\sigma}^2_T) \xrightarrow{p} 0$. The variance can be expressed as

$$\begin{aligned} &E[\sqrt{T}(\hat{\sigma}^2_{m,T} - m\hat{\sigma}^2_T)]^2 = TE[(\hat{\sigma}^2_{m,T} - m\sigma^2) - m(\hat{\sigma}^2_T - \sigma^2)]^2 \\ &\quad = T[E(\hat{\sigma}^2_{m,T} - m\sigma^2)^2 + m^2 E(\hat{\sigma}^2_T - \sigma^2)^2 - 2mE(\hat{\sigma}^2_{m,T} - m\sigma^2)(\hat{\sigma}^2_T - \sigma^2)]. \end{aligned} \tag{A13}$$

As before we have the established result $\hat{\sigma}^2_T = (1/T)\hat{\boldsymbol{u}}'_T\hat{\boldsymbol{u}}_T \xrightarrow{p} (1/T)\boldsymbol{u}'_T\boldsymbol{u}_T \xrightarrow{p} \sigma^2$ and $\sqrt{T}(\hat{\sigma}^2_T - \sigma^2) \xrightarrow{d} N(0,(\mu_4 - \sigma^4))$.

For $\hat{\sigma}^2_{m,T}$, with reference to model (3) we re-define the earlier aggregation matrix $\boldsymbol{A}$ in (A2) to be of dimension $(T-m+1)\times T$ with θ replaced with ρ. Further, redefine $\boldsymbol{\beta}^*$ as $\boldsymbol{\beta}^* = (\rho, \beta, \phi_1, ..., \phi_p)'$, $\boldsymbol{\beta}$ is the same as before, replace the first element of *t*th row of $\boldsymbol{X}^*$ with y_{t-1}, and $\boldsymbol{V} = \boldsymbol{A}\boldsymbol{u}$, ($V_t = u_t + \rho u_{t-1} + ... + \rho^{m-1}u_{t-m+1}$). Again using the vector-matrix notation similar to that of Theorem 1 we obtain under the null $\rho = 1$:

$$
\begin{aligned}
\hat{\sigma}^2_{m,T} &= (1/T)\hat{\boldsymbol{V}}_T'\hat{\boldsymbol{V}}_T = (1/T)\boldsymbol{V}_T'\boldsymbol{V}_T - (2/T)(\hat{\boldsymbol{\beta}}^*_T - \boldsymbol{\beta}^*)'\boldsymbol{X}^{*\prime}_{aT}\boldsymbol{V}_T \\
&\qquad + (\hat{\boldsymbol{\beta}}^*_T - \boldsymbol{\beta}^*)'(\boldsymbol{X}^{*\prime}_{aT}\boldsymbol{X}^*_{aT}/T)(\hat{\boldsymbol{\beta}}^*_T - \boldsymbol{\beta}^*) \\
&= (1/T)\boldsymbol{V}_T'\boldsymbol{V}_T - 2(\hat{\boldsymbol{\beta}}^*_T - \boldsymbol{\beta}^*)'(\Upsilon_T/\sqrt{T})(\Upsilon_T^{-1}/\sqrt{T})\boldsymbol{X}^{*\prime}_{aT}\boldsymbol{V}_T \qquad \text{(A14)} \\
&\quad + (\hat{\boldsymbol{\beta}}^*_T - \boldsymbol{\beta}^*)'(\Upsilon_T/\sqrt{T})(\Upsilon_T^{-1}\boldsymbol{X}^{*\prime}_{aT}\boldsymbol{X}^*_{aT}\Upsilon_T^{-1})(\Upsilon_T/\sqrt{T})(\hat{\boldsymbol{\beta}}^*_T - \boldsymbol{\beta}^*) \\
&\xrightarrow{p} m\sigma^2.
\end{aligned}
$$

This result holds because $(\Upsilon_T/\sqrt{T})(\hat{\boldsymbol{\beta}}^*_T - \boldsymbol{\beta}^*) \xrightarrow{p} \boldsymbol{0}$ while the rest converge to bounded quantities.

Now we have to consider the distribution of $\sqrt{T}(\hat{\sigma}^2_{m,T} - m\sigma^2)$. Multiplying (A14) through by $\sqrt{T}$ shows that the last term of (A14) converges in probability zero and in the second term, $T(\hat{\rho}-1)(T^{-3/2}\sum X_{a1t}V_t) \xrightarrow{p} 0$ and $\sqrt{T}(\hat{\beta}-\beta)(1/T)\sum X_{a2t}V_t \xrightarrow{p} 0$. Thus we have to consider the distribution of :

$$
\sqrt{T}(\hat{\sigma}^2_{m,T} - m\sigma^2) = \sqrt{T}(\boldsymbol{V}_T'\boldsymbol{V}_T/T_a - m\sigma^2) - 2\sqrt{T}(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta})'(\boldsymbol{X}'_{aT}\boldsymbol{V}_T/T_a) \qquad \text{(A15)}
$$

where $T_a = T - m + 1$.

Now consider the first term on the RHS of (A15). Its variance can be written as

$$
\begin{aligned}
& E[(1/\sqrt{T})(\sum(V_t^2 - m\sigma^2)]^2 \\
&\quad = (1/T)[\sum E(V_t^2 - m\sigma^2)^2 + 2\sum_{k>0} E(V_t^2 - m\sigma^2)(V_{t-k}^2 - m\sigma^2) \\
&\quad = \{\gamma_0 + (2/T)[(T-1)\gamma_1 + (T-2)\gamma_2 + \ldots + (T-m+1)\gamma_{m-1}]\} \qquad \text{(A16)} \\
&\quad = \gamma_0 + 2\sum_{k=1}^{m-1}\gamma_k - (2/T)\sum_{k=1}^{m-1}k\gamma_k
\end{aligned}
$$

where γ represents the variance and covariance terms and as we shall see later $\gamma_k = 0$, for $k \geq m$.

Now consider the variance term in (A16):

$$\gamma_0 = E(V_t^2 - m\sigma^2)^2 = E(V_t^4 + m^2\sigma^4 - 2m\sigma^2 V_t^2)$$
$$= E(V_t^4) - m^2\sigma^4$$

where

$$E(V_t^4) = E(\sum_{i=0}^{m-1} u_{t-i})^4 = E[(\sum_{i=0}^{m-1} u_{t-i})^2]^2$$
$$= \mathrm{E}[\sum u_{t-i}^2 + 2\sum_{i<j} u_{t-i}u_{t-j}]^2$$
$$= \mathrm{E}[(\sum u_{t-i}^2)^2 + 4(\sum_{i<j} u_{t-i}u_{t-j})^2 + 4(\sum u_{t-i}^2)(\sum_{i<j} u_{t-i}u_{t-j})]$$
$$= \sum E(u_{t-i}^4) + 2\sum_{i<j} E(u_{t-i}^2 u_{t-j}^2) + 4\sum_{i<j} E(u_{t-i}^2 u_{t-j}^2)$$
$$= m\mu_4 + 6\frac{m(m-1)}{2}\sigma^4$$
$$= m\mu_4 + 3m(m-1)\sigma^4.$$

In this expression, the terms that become zero upon taking expectation have been dropped out. Combining the terms we get

$$\gamma_0 = E(V_t^2 - m\sigma^2)^2 = m\mu_4 + 3m(m-1)\sigma^4 - m^2\sigma^4$$
$$= m\mu_4 + 2m^2\sigma^4 - 3m\sigma^4 \qquad \text{(A17)}$$

Now consider the covariance terms in (A16):

$$E(V_t^2 - m\sigma^2)(V_{t-k}^2 - m\sigma^2) = E(V_t^2 V_{t-k}^2) - m^2\sigma^4$$

where

$$
\begin{aligned}
E(V_t^2 V_{t-k}^2) &= E[(\sum_{i=0}^{m-1} u_{t-i})^2 (\sum_{i=0}^{m-1} u_{t-k-i})^2] \\
&= E[(\sum u_{t-i}^2 + 2\sum_{i<j} u_{t-i}u_{t-j})(\sum u_{t-k-i}^2 + 2\sum_{i<j} u_{t-k-i}u_{t-k-j})] \\
&= E(\sum u_{t-i}^2 \sum u_{t-k-i}^2) + 4E(\sum_{i<j} u_{t-i}u_{t-j} \sum_{i<j} u_{t-k-i}u_{t-k-j}).
\end{aligned}
$$

In this expression, the two cross product terms that become zero upon taking expectation have been dropped out. By evaluating the above two terms we obtain:

$$
\begin{aligned}
&\gamma_1 = (m-1)\mu_4 + (m^2 - (m-1))\sigma^4 + (1/2)(m-1)(m-2)\sigma^4 - m^2\sigma^4 \\
&\gamma_2 = (m-2)\mu_4 + (m^2 - (m-2))\sigma^4 + (1/2)(m-2)(m-3)\sigma^4 - m^2\sigma^4 \\
&. \\
&\gamma_{m-1} = \mu_4 + (m^2 - 1))\sigma^4 - m^2\sigma^4 \\
&\gamma_k = m^2\sigma^4 - m^2\sigma^4 = 0,\ k \geq m.
\end{aligned}
$$

From these we obtain:

$$
\begin{aligned}
\sum_{k=1}^{m-1}\gamma_k &= \mu_4\sum_{k=1}^{m-1}(m-k) + [\sum_{k=1}^{m-1}(m^2 - (m-k)) + (1/2)\sum_{k=1}^{m-1}(m-k)(m-k-1) \\
&\quad - m^2(m-1)]\sigma^4 \qquad \text{(A18)} \\
&= (1/2)m(m-1)\mu_4 + (1/6)m(m-1)(4m-11)\sigma^4
\end{aligned}
$$

and

$$
\begin{aligned}
\sum_{k=1}^{m-1}k\gamma_k &= \mu_4\sum_{k=1}^{m-1}k(m-k) + [\sum_{k=1}^{m-1}k(m^2 - (m-k)) + (1/2)\sum_{k=1}^{m-1}k(m-k)(m-k-1) \\
&\quad - m^2 m(m-1)/2]\sigma^4 \\
&= (1/2)m(m-1)^2\mu_4 + (1/24)m(m^2-1)(m-6)\sigma^4
\end{aligned}
$$

(A19)

In our simulation results we observe that the term in (A19) does not make much contribution to the results in small samples and therefore we simply use the asymptotic result. Therefore, adding (A17) and 2 times (A18) yields the variance of the first term on the RHS of (A15):

$$E[(1/\sqrt{T})\sum(V_t^2 - m\sigma^2]^2 = m^2\mu_4 + m(4m-1)(m-2)\sigma^4/3. \qquad \text{(A20)}$$

Now consider the second term on the RHS of (A15). To obtain its variance first note that

$\sqrt{T}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow{d} N(\boldsymbol{0}, \sigma^2(\boldsymbol{X}'\boldsymbol{X}/T)^{-1})$ and $\boldsymbol{X}_{aT}{}'\boldsymbol{V}_T/T \xrightarrow{p} \boldsymbol{c}(\sigma^2, \phi_1, ..., \phi_p, m)$

where $\boldsymbol{c}$ is a $(1+p)\times 1$ vector. This vector can be derived easily by noting that the aggregated form of model (3) under the null provides, $\Delta_m y_t = \phi(L)^{-1}V_t = \psi(L)S(L)u_t$, where $\phi(L)^{-1} = \psi(L) = 1 + \psi_1 L + \psi_2 L^2 + ...$ and $S(L) = 1 + L + L^2 + ... + L^{m-1}$.

Note that the first term of $\boldsymbol{X}_{aT}{}'\boldsymbol{V}_T/T_a$ that corresponds to the constant term of the model is zero. Now consider the second term in the $\boldsymbol{X}_{aT}{}'\boldsymbol{V}_T/T_a$ vector:

$$\begin{aligned}(1/T)\sum X_{a2t}^* V_t &= (1/T)\sum \Delta_m y_{t-1} V_t = (1/T)\sum(\psi(L)S(L)u_{t-1})S(L)u_t \\ &= (1/T)\sum[(1 + (1+\psi_1)L + (1+\psi_1+\psi_2)L^2 + ...)u_{t-1}](1 + L + L^2 + ... + L^{m-1})u_t \\ &= \sigma^2[(m-1) + (m-2)\psi_1 + (m-3)\psi_2 + ... + \psi_{m-2}].\end{aligned}$$

Proceeding in this way, we obtain for *p=m-1*:

$$\boldsymbol{c} = \sigma^2 \begin{bmatrix} 0 \\ (m-1)+(m-2)\psi_1+(m-3)\psi_2+\ldots+\psi_{m-2} \\ (m-2)+(m-3)\psi_1+(m-4)\psi_2+\ldots+\psi_{m-3} \\ . \\ . \\ 1 \end{bmatrix}. \tag{A21}$$

The $\boldsymbol{c}$ vector can easily be computed from $(1/T)\sum \Delta_m y_{t-k}\hat{V}_t$, k=1,2,…,p. If $p \geq m$ the $\boldsymbol{c}$ vector will have zero entries for the excess terms. Using these results the variance of the second term of (A15) can be written as

$$Var[(2/\sqrt{T})(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta})'\boldsymbol{X}_a'\boldsymbol{V}] = 4\sigma^2 \boldsymbol{c}'(\boldsymbol{X}'\boldsymbol{X}/T)^{-1}\boldsymbol{c}. \tag{A22}$$

Now using Hausman's approach (Hausman, 1978) the required variance can be obtained as the difference of the asymptotic variances. Using (A20) and (A22) we have:

$$\begin{aligned} &Var[(1/\sqrt{T})(\boldsymbol{V_T'V_T} - Tm\sigma^2) - (2/\sqrt{T})(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta})'\boldsymbol{X}_a'\boldsymbol{V}] \\ &\quad = (m^2\mu_4 + m(4m-1)(m-2)\sigma^4/3) - 4\sigma^2\boldsymbol{c}'(\boldsymbol{X}'\boldsymbol{X}/T)^{-1}\boldsymbol{c}. \end{aligned} \tag{A23}$$

Note that $(V_t^2 - m\sigma^2)$ is a stationary process and therefore by the central limit theorem

$$\begin{aligned} &\sqrt{T}(\hat{\sigma}^2_{m,T} - m\sigma^2) \xrightarrow{d} \\ &\quad N(0,\ m^2\mu_4 + m(4m-1)(m-2)\sigma^4/3 - 4\sigma^2\boldsymbol{c}'(\boldsymbol{X}'\boldsymbol{X}/T)^{-1}\boldsymbol{c}) \end{aligned}. \tag{A24}$$

Although the covariance term of (A13) can be worked out we can apply the Hausman approach again to obtain the overall variance of $\sqrt{T}(\hat{\sigma}^2_{m,T} - m\hat{\sigma}^2)$:

$$E[\sqrt{T}(\hat{\sigma}_m^2 - m\hat{\sigma}^2)]^2 = 2m(2m-1)(m-1)\sigma^4/3 - 4\sigma^2 \boldsymbol{c}'(\Upsilon^{-1}\boldsymbol{X}'\boldsymbol{X}\Upsilon^{-1})^{-1}\boldsymbol{c}\,. \tag{A25}$$

This is the variance of the sum of two asymptotically normally distributed variables, hence we establish that

$$\sqrt{T}(\hat{\sigma}_{m,T}^2 - m\hat{\sigma}_T^2) \xrightarrow{d} N(0, 2m(2m-1)(m-1)\sigma^4/3 - 4\sigma^2 \boldsymbol{c}'(\Upsilon^{-1}\boldsymbol{X}'\boldsymbol{X}\Upsilon^{-1})^{-1}\boldsymbol{c}). \tag{A26}$$

QED

Through simulations in small samples we observe that the degrees of freedom in estimating $\hat{\sigma}_m^2$ in (A26) plays an important role in obtaining the correct size of the test. To derive the degrees of freedom express the sum of squares of $\hat{V}_t$ in matrix form as $\sum_{t=m}^{T} \hat{V}_t^2 = \hat{\boldsymbol{V}}'\hat{\boldsymbol{V}}$. By taking expectation of this we get under the null

$$\begin{aligned} E(\hat{\boldsymbol{V}}'\boldsymbol{V} \mid \boldsymbol{X}) &= E[(\boldsymbol{AMu})'(\boldsymbol{AMu}) \mid \boldsymbol{X}] \\ &= E[tr(\boldsymbol{AMu})'(\boldsymbol{AMu}) \mid \boldsymbol{X}] = \sigma^2 tr(\boldsymbol{AMA}') \\ &= \sigma^2[tr(\boldsymbol{AA}') - tr(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}_a'\boldsymbol{X}_a] \\ &\approx \sigma^2[m(T-m+1) - 2m^2(T-m+1)/T \\ &= m\sigma^2(T-m+1)(1-2m/T). \end{aligned}$$

This approximation holds exactly for p=0 in (3). If a constant term is not included in the model *2m* has to be replaced with *m*. Here we have used $\boldsymbol{M} = \boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$, where $\boldsymbol{X}$ represents the full design matrix from (4).

References

Abeysinghe, T, (1998) "Forecasting Singapore's quarterly GDP with monthly external trade", *International Journal of Forecasting*, 14, 505-513.

Abeysinghe, T, (2000) "Modeling variables of different frequencies", *International Journal of Forecasting*, 16, 117-119

Abeysinghe, T. and Choy, K.M. (2007) *The Singapore Economy: An Econometric Perspective*, Routledge, London.

Bhargava, A. (1986) "On the theory of testing for unit roots in observed time series". *Review of Economic Studies*, 53, 369-384.

Breidt, F.J., Davis, R.A., Nan-Jung Hsu and Rosenblatt, M. (2006) "Pile-up probabilities for the Laplace likelihood estimator of a non-invertible first order moving average", *IMS Lecture Notes (Monograph Series)*, Institute of Mathematical Statistics.

Breitung J. (1994) "Some simple tests of the moving-average unit root hypothesis", *Journal of time series analysis*, 15, 351-370.

Choi, I. (1992) "Durbin-Hausman Tests for a Unit Root" *Oxford Bulletin of Economics and Statistics*, 54, 289-304.

Choi I. (1994) "Residual-based tests for the null of stationary with application to US Macroeconomic time series", *Econometric Theory*, 10, 720-746.

Davis, R. and Dunsmuir, W. (1996) "Maximum likelihood estimation for MA(1) processes with a root on or near the unit circle", *Econometric Theory*, 12, 1-29.

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

DeJong, D.N., Nankervis, J.C., Savin, N.E., and Whiteman, C.H. (1992) "Integration versus trend stationary in time series", *Econometrica*, 60, 423-433.

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

Harvey, A. C. (1989) *Forecasting Structural Time Series Models and the Kalman Filter*, Cambridge: Cambridge University Press.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. and Shin, Y. (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.

Lee, J. and Schmidt, P. (1994) “Unit root tests based on instrumental variables estimation”, *International Economic Review,* 35, 449-462.

Leybourne S.J. and B.P.M. McCabe (1994) “A consistent test for a unit root”, *Journal of business and economic statistics,* 12, 157-166.

Leybourne, S.J. (1995) “Testing for Unit Roots Using Forward and Reverse Dickey-Fuller Regressions” *Oxford Bulletin of Economics and Statistics*, 57, 559-571.

Lo A.W. and MacKinlay, A.C. (1988) “Stock prices do not follow random walks: Evidence from a simple specification test”, *The Review of Financial Studies,* 1, 41-66.

Lo A.W. and MacKinlay, A.C. (1989) “The size and power of the variance ratio test in finite samples”, *Journal of Econometrics,* 40, 203-238.

Maddala G.S. and Kim, I. (1998), *Unit roots, cointegration and structural change,* Cambridge: Cambridge University Press.

Pantula S.G. and A. Hall (1991) “Testing for unit roots in autoregressive moving average models” *Journal of econometrics,* 48, 325-353

Pantula, S.G., Gonzalez-Farias, G. and Fuller, W.A. (1994) “A Comparison of Unit-Root Test Criteria”, *Journal of Business and Economic Statistics,* 12, 449-459.

Park J.(1990) “Testing for unit roots and cointegration by variable addition”, *Advances in Econometrics,* 8, 107-133.

Park, H.J. and W.A. Fuller (1995) “Alternative Estimators and Unit Root Tests for the Autoregressive Process”, *Journal of Time Series Analysis,* 16, 415-429.

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.

Phillips, P.C.B. and Perron, P. (1988) “ Testing for a unit root in time series regression”, *Biometrica,* 75, 335-346.

Phillips, P.C.B. (1998) "New Tools for Understanding Spurious Regressions." *Econometrica,* 66, 1299-1325.

Phillips, P.C.B. (2002) "New Unit Root Asymptotics in the Presence of Deterministic Trends," *Journal of Econometrics,* 111, 323-353.

Rajaguru, G. (2004) *Effects of temporal aggregation and systematic sampling on model dynamics and causal inference*, PhD Thesis, Department of Economics, National University of Singapore.

Rajaguru. G. Abeysinghe, T. (2002) The distortionary effects of temporal aggregation on Granger causality" Working Paper wp0204, Department of Economics, National University of Singapore.

Saikkonen P. and R. Luukkonen (1993) "Testing for a moving average unit root in ARIMA models", *Journal of American Statistical Association*, 88, 596-601.

Schmidt, P. and P.C.B. Phillips (1992) "LM tests for a unit root in the presence of deterministic trends" *Oxford Bulletin of Economics and Statistics*, 54, 257-287.

Blough, S.R. (1992) "The relationship between power and level for Generic Unit Root tests in finite samples" *Journal of applied econometrics*, 7, 295-308.

Yap S.F. and Reinsel, G.C. (1995) "Results on estimation and testing for a unit root in the non-stationary ARMA model", *Journal of time series analysis*, 16, 339-353.

Table 1: Power of unit root tests at the 5% level and T=100

Reference model: $y_t = \alpha + \beta t + \rho y_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1},\ \varepsilon_t \sim iid(0,\sigma^2)$

(When T=100 is not available 200 is used and marked with an asterisk against author's name)

(a) Non-stationary null (ρ = 1)

Name of Authors	**Year**	**Model Type**	**Test Type**	**ρ = 0.80**	**0.85**	**0.90**	**0.95**	**0.975**	**Remarks**
Dicky & Fuller	1979	θ=0, β=0	$\hat{\rho}$	0.86		0.30	0.10		
		θ=0, β=0	t	0.73		0.18	0.06		DF test, AR(1) process
Bhargava	1986	θ=0, β=0	DW	0.73	0.49	0.25	0.10		Also Sargan & Bhargava 1983
Phillips & Perron	1988	θ=0, β=0	t		0.47				ADF, Said & Dicky 1984
		θ=0.8, β=0	t		0.30				ADF
		θ=0, β=0	Z(t)		0.69				PP
		θ=0.8, β=0	Z(t)		0.35				PP
Pantula & Hall*	1991	θ=0, β=0	IV					0.09-0.33	Range of IV estimates. In
		θ=0.8, β=0	IV					0.01-0.35	general power > 0.05
DeJong et al.	1992	θ=0, β≠0	τ(ρ)	0.75	0.49	0.24	0.10		For starting value 0. Power drops slightly as starting value
			F(β,ρ)	0.65	0.39	0.19	0.08		increases.
Blough	1992	θ=0, β=0	ADF, IV						Graphical presentation. Power drops to 5% for ρ>0.5.
Schmidt & Phillips	1992	θ=0, β≠0	LM			0.27	0.108		Reported is highest power under different specifications
Choi	1992	θ=0, β≠0	DH	0.97	0.84	0.54	0.24		Durbin-Hausman
Lee & Schimidt	1994	θ=0.8, β=0	IV				0.22		Compares Hall-IV with SP-IV
Pantula et al.	1994	θ=0, β=0	WS			0.602	0.261		Compares OLS, MLE as well.
Yap & Reinsel *	1995	θ=0, β=0	LR	1.00		0.82	0.33		
		θ=0.8, β=0	LR	-		0.74	0.56		
Leybourne	1995	θ=0, β=0	DFmax	0.88		0.34			

Table 1 continued

Name of Authors	**Year**	**Model Type**	**Test Type**	**ρ = 0.80**	**0.85**	**0.90**	**0.95**	**0.975**	**Remarks**
Park & Fuller	1995	θ=0, β=0							Graphical. For intercept model: WS>SS>OLS. For interceptless model: OLS>SS>WS. (SS=simple symmetric, WS=weighted symmentric)
Perron & Ng *	1996	θ=0.8, β=0	MZ(ρ)			0.75	0.42		
			MSB			0.79	0.46		
			MZ(t)			0.63	0.30		Modified PP
Elliot et al.	1996	θ=0.8, β=0	t	0.51		0.30	0.15		Power at ρ=0.95 not very different across models
Hwang & Schmidt	1996	θ=0, $\beta\neq$0	GLS	0.28	0.18				Power is roughly similar across different tests reported

Non-stationary null: Structural breaks

Lanne & Lutkepohl	2002	Perron				0.21			Known break, level shift. Power is very similar for slope change. See the article for model specification.
		Perron & Vogelsang				0.14			
		Amsler & Lee				0.12			
		Schmidt & Phillips				0.09			
		Lanne et al				0.23			
Lanne et al.	2003	Test 1, drift				0.28			Unknown break, level shift. Power is very similar for slope change. See the article for model specification.
		Test 2, drift				0.20			
		Test 3, trend				0.23			
		Test 3, trend				0.18			

Table 1 continued

(b) Stationary null ($\rho = 1$, $\theta = 1$)

Name of Authors	**Year**	**Model Type**	**Test Type**	**θ = 0.80**	**0.85**	**0.90**	**0.95**	**0.975**	**Remarks**
Park	1990		J1 test						No simulation results
Kwiatkowski et al.	1992	β=0	$\eta(\mu)$ l0			0.59		0.17	
		β=0	$\eta(\mu)$ l4			0.51		0.15	
		β=0	$\eta(\mu)$ l12			0.38		0.10	
		$\beta \neq 0$	$\eta(\tau)$ l0			0.35		0.05	KPSS test. The test basically
		$\beta \neq 0$	$\eta(\tau)$ l4			0.28		0.05	involves testing σ_η^2 = 0 in
		$\beta \neq 0$	$\eta(\tau)$ l12			0.17		0.04	model (1) in Section 3.
Saikkonen & Luukkonen	1993	β=0	R2	0.81	0.71	0.56	0.32		Authors also consider non-white errors.
Breitung	1994	β=0	Spectral	0.04		0.03	0.03		
			Var diff	0.87		0.43	0.16		
			Tanaka	0.86		0.62	0.32		
Leybourne and McCabe	1994	Extended KPSS	s(α) p=1			0.61		0.17	Show that KPSS is subject to
			s(α) p=2			0.59		0.17	severe size distortions in
			s(α) p=3			0.56		0.16	general ARIMA cases.
Choi	1994	β=0	w1 l=2	0.47					
			w1 l=3	0.38					
			w1 l=4	0.27					
			w1 l=5	0.06					Power remains low for other
		$\beta \neq 0$	w2 l=1	0.08					lags on w2 test