Working Paper 92-7/Document de travail 92-7

Unit Root Tests and the Burden of Proof

by
Robert A. Amano and Simon van Norden

Bank of Canada Banque du Canada

Unit-Root Tests and the Burden of Proof

Robert A. Amano and Simon van Norden*
International Department
Bank of Canada
Ottawa, Ontario K1A 0G9

November 1992

* We thank Norm Cameron, Pat Durnford, Serena Ng, Wayne Simpson and seminar participants at the 1992 Canadian Economic Association meetings for their helpful comments. The responsibility for errors is ours. Opinions expressed here do not necessarily reflect those of the Bank of Canada or its staff.

Abstract

Simulation evidence is presented on the finite sample properties of two tests for stationarity recently proposed by Kwiatkowski, Phillips and Schmidt (1991) and Park (1990). Unlike earlier unit-root tests, these test the null of stationarity against the alternative of a unit root, thereby reversing the usual burden of proof. We also examine the consequences of using the Kwiatkowski, Phillips and Schmidt test in conjunction with a standard unit-root test. These results suggest that the frequency of incorrect conclusions may be decreased relative to the application of only standard unit-root tests. Also, such a joint testing procedure may in some cases permit researchers to be more confident about their tests' results.

Résumé

Les auteurs présentent les résultats des simulations qu'ils ont effectuées afin de vérifier les propriétés, avec des échantillons finis, des tests de stationnarité récemment proposés par Kwiatkowski, Phillips et Schmidt (1991) et par Park (1990). Au contraire des tests usuels de racine unitaire, ces deux approches présentent le cas de stationnarité comme l'hypothèse nulle à rejeter lorsque celle-ci est confrontée à l'hypothèse alternative de racine unitaire. Les auteurs examinent également les conséquences d'une utilisation conjointe du test de Kwiatkowski, Phillips et Schmidt et des tests types de racine unitaire. Les résultats indiquent que, avec cette approche, il est possible d'en arriver moins souvent à des conclusions erronées que si l'on n'applique que des tests usuels de racine unitaire. Par conséquent, cette approche conjointe peut dans certains cas accroître la confiance des chercheurs dans les résultats de leurs tests.

1.0 Introduction

Standard unit-root tests such as those of Dickey and Fuller (1979) and Phillips and Perron (1988) frequently do not reject the null of a unit root when they are applied to macroeconomic time series.[1] The presence of a unit root raises serious questions both about the existence and nature of the business cycle and about the econometric methods used to draw inferences from macroeconomic data. However, many researchers have pointed out that unit-root tests have little power against stationary alternatives with high persistence and argue that the presence of a unit root is therefore unproven. For example, DeJong, Nankervis, Savin and Whiteman (1992) show that the Dickey and Fuller test is typically unable to distinguish between series with a unit root and stationary AR(1) series with an autocorrelation coefficient near unity. An alternative way to approach the question would be to use Bayesian methods to examine the probability that a unit root is consistent with the data.[2] The results have generated considerable debate, however, and seem to be sensitive to normally innocuous assumptions about prior beliefs.[3] We instead focus on a third approach, based on procedures designed to test the null of stationarity against the alternative of a unit root, thereby reversing the usual burden of proof.

Recent papers by Kwiatkowski, Phillips and Schmidt (1991) and Park (1990) have proposed different tests of the null hypothesis of stationarity against the alternative of a unit root.[4] However, little is known about the finite sample behaviour of these new tests for stationarity, their robustness and their relative strengths and weaknesses. We attempt to address these questions by using Monte Carlo simulations to calculate the empirical size and power of these tests for selected data generation processes. We also examine the consequences of using these tests in conjunction with a standard unit-root test. By comparing the results of such tests to standard unit-root tests, one

---

1. For example, see Abuaf and Jorion (1983) for evidence on real exchange rates, Diba and Grossman (1988) for the price-dividend ratio, Nelson and Plosser (1982) for output, and Rose (1990) for real interest rates.
2. For an introduction see Sims (1988).
3. See Phillips (1991a,1991b).
4. Another test for stationarity has recently been proposed by Kahn and Ogaki (1992). However, the test's applicability appears limited, as it does not extend to cases where the residuals are not white noise. Dejong, Nankervis, Savin and Whiteman (1992) propose both similar and non-similar tests for the null of stationarity. Unfortunately, their similar tests consider only the null of a specific AR root rather than a general null of stationarity, and their non-similar test will generally lack power due to nuisance parameter problems.

can hope to determine whether there is significant evidence of a unit root, or whether the data are simply uninformative. We find that in some cases the joint test helps to avoid erroneous conclusions. A simple Bayesian analysis suggests that for certain priors, the joint testing procedure will allow researchers to be more confident about their tests' results when the two tests agree. It also introduces a significant probability that the two tests will fail to agree in their results.

The next section briefly introduces the Kwiatkowski, Phillips and Schmidt (KPS) and Park tests. Section 3 presents our Monte Carlo design and the results on test size, power and robustness. Section 4 investigates the usefulness of applying the KPS test in conjunction with an existing unit-root test. Section 5 provides concluding remarks.

2.0 Two Tests for the Null Hypothesis of Stationarity[5]

The Dickey and Fuller (DF) and Phillips and Perron (PP) tests for unit roots are based on a test regression of the form

$$y_t = \alpha + \rho y_{t-1} + v_t, \tag{1}$$

where v_t is assumed to be stationary.[6] KPS propose a slightly different testing framework based on the model

$$y_t = r_t + \varepsilon_t, \tag{2}$$

where

$$r_t = r_{t-1} + u_t, \tag{3}$$

u_t is *i.i.d.* and ε_t is assumed to be stationary. Equation (2) decomposes y_t into a random walk component r_t and a stationary error. The initial value of the random walk component is treated as fixed and so serves as an intercept or the mean to which the series reverts. The null hypothesis of

5. We consider only tests comparing hypotheses of a unit root without drift to stationary hypotheses. Extensions to the case of a unit root with drift and trend stationarity are straightforward.
6. The tests differ in their treatment of serially correlated residuals. DF propose adding lagged differences to control for serial correlation, while PP propose the application of a nonparametric correction to the test statistic. In fact, the PP correction allows for more general dependence in the residual process, including conditional heteroscedasticity.

stationarity is $\sigma_u^2 = 0$ or, equivalently, $\sigma_r^2 = 0$.

The key insight of KPS is that this is simply a special case of the random coefficients model

$$y_t = x_t \beta_t + \varepsilon_t, \tag{4}$$

where

$$\beta_t = \beta_{t-1} + u_t, \tag{5}$$

where imposing $x_t = 1$ and $\beta_t = r_t$ reduces Equation (4) to (2). This means that testing the null of stationarity against the alternative of a unit root is equivalent to testing the null of a constant coefficient model against the alternative of a random coefficient model. The appropriate Lagrange Multiplier test statistic works out to be

$$\text{LM} = \frac{\sum_{t=1}^{T} S_t^2}{\hat{\sigma}_\varepsilon^2}, \tag{6}$$

where

$$S_t = \sum_{i=1}^{t} e_i \tag{7}$$

and e_t is the residual from the regression of y_t on a constant, $\hat{\sigma}^2_\varepsilon$ is the usual estimate of the residual variance from this regression (residual sum of squared residuals divided by *T-1*), and *T* is the sample size. The distribution of this statistic is nonstandard and empirical critical values are provided by Monte Carlo methods.

In the random coefficients literature, this derivation depends on the strong assumption that ε_t is *i.i.d.* normal with mean zero. KPS extend this test to the case where ε_t is serially correlated and may be heteroscedastic.[7] The key difference is that the denominator of (6) is replaced with the Newey and West (1987) estimate of the long-run variance of ε_t, denoted $s^2(k)$, giving the KPS

7. Any stationary ARMA process is acceptable. See KPS for a precise statement of the conditions ε_t must satisfy.

test statistic[8]

$$\hat{\eta} = \frac{T^{-2} \cdot \sum_{t=1}^{T} S_t^2}{s^2(k)}, \tag{8}$$

where k is equal to the lag truncation parameter and an additional factor T^{-2} is used to normalize the numerator.

An alternative test proposed by Park (1990) is based on the idea that if a variable follows a unit-root process, then ordinary least-squares standard errors are usually inappropriate and will tend to indicate that unrelated variables have a statistically significant relationship. Park's J1 test adds one or more spurious variables to a regression of y_t on a constant and tests whether they appear to be significant. If we know a priori that these regressors are superfluous, then their nominal significance in standard F- or t-tests is an indication of a spurious regression. The Park J1 test statistic is of the form

$$J1 = \frac{RSS_1 - RSS_2}{\omega^2(k)}, \tag{9}$$

where RSS_1 is the sum of y_t's squared deviations from its mean, RSS_2 is the residual sum of squares of the regression with a constant and the superfluous variables, and $\omega^2(k)$ is the Newey and West estimate of the long-run residual variance of y, defined analogously to $s^2(k)$ above. Under the null hypothesis of stationarity, the test statistic is Chi-squared distributed with degrees of freedom equal to the number of superfluous regressors. The selection of the superfluous regressors is somewhat arbitrary, with any deterministic trend or unrelated stochastic process a legitimate choice. Park (1990) suggests polynomial time trends and pseudo-random walks. Park, Ouliaris and Choi (1988) note that in finite samples the power of the test appears to vary greatly with the number of superfluous regressors, although the selection of the superfluous variables quickly becomes unimportant as the sample size increases. They also found that using two or more variables

8. Newey and West define $s^2(k) \equiv T^{-1} \cdot \left[\sum_{t=1}^{T} e_t^2 + 2\sum_{s=1}^{k} w(s,k) \sum_{t=s+1}^{T} e_t e_{t-s}\right]$ where $w(s,k) = 1 - s/(k+1)$, which ensures that the estimator is non-negative. Note that when k = 0, the estimator is simply the standard variance estimator.

generally gave better discriminatory power. In this paper, we include three time trends (linear, cubic and fifth order) as superfluous regressors, as this case is thought to have good power and is more useful to the applied researcher than that using pseudo-random walks.[9]

3.0 Monte Carlo Experiment Design and Results

In this section, we use Monte Carlo experiments to evaluate and compare the effects of different data generation processes (DGPs) on the size and power of the KPS and J1 tests. The simulated data are generated by an ARMA(1,1) model of the form

$$y_t = \rho y_{t-1} + \vartheta_t + \theta \vartheta_{t-1}, \tag{10}$$

where ϑ_t is *i.i.d.*, with mean zero and constant variance equal to one and $|\theta| < 1$. For each experiment we perform 10,000 replications with the sample size T set equal to 50, 100 and 200 observations, and we record the rejection frequencies using 5 per cent asymptotic critical values.[10] Various lag truncation parameters (0, 1, 2, 4 and 8) are used to evaluate the performance of the tests.[11] To estimate the size of the tests, we set $\rho = 0.85$ for a range of MA parameters (-0.8, -0.5, 0.0, 0.5 and 0.8), while power is calculated by setting $\rho = 1.0$ for the same range of MA parameters. For sample sizes usually available in macroeconomics, this set of AR parameters is typical of the range in which the parameters of interest are thought to lie. DeJong, Nankervis, Savin and Whiteman (1992) present Monte Carlo results to show that the most commonly used unit-root tests can not reliably distinguish between such processes for typical sample sizes. Therefore, this is precisely the area of the parameter space in which better tests are most needed. Furthermore, restricting the number of AR parameters examined allows us to keep the number of permutations

9. Results using three pseudo-random walks are available from the authors on request.
10. The 5 per cent critical values are 0.463 and 7.815 for the KPS and J1 tests respectively. We begin data generation at T= -19 and discard the first 20 observations to minimize any problems associated with starting values.
11. All simulations were performed on a Sun Sparcstation using RATS Version 3.1.

examined down to a manageable level.[12] Finally, we note that our results for varying sample sizes can be interpreted either as cases where the span of the data varies for a given sampling frequency, or where the sampling frequency varies for a given data span. In the latter case, holding ρ constant at 0.85 while varying *T* implies we are changing the adjustment speed of the series. Therefore, results across different values of *T* can give insight into the test's behaviour across different values of ρ.[13]

Table 1 (p. 12) compares size for the KPS and J1 tests. For each lag length *k*, two rejection frequencies are reported; the first corresponds to the KPS test and the second to the J1 test. In general, both tests show large size distortions that tend to increase with the value of the MA parameter and decrease with the lag truncation length. The J1 test appears to be slightly more conservative than the KPS test across most DGPs and sample sizes. However, for sample size equal to 50 and truncation parameter equal to 8, the J1 test is too conservative, with size calculations less than the nominal size of 5 per cent. A striking feature of Table 1 is that for shorter truncation lags, the size of the tests tends to move in opposite directions (for a given *k*) over different sample sizes. That is, size distortion corresponding to the J1 test tends to increase for larger sample sizes, while that for the KPS test tends to decrease for larger samples. However, for lag lengths greater than 2, size distortions tend to move in the same direction.[14]

Unadjusted power calculations appear in Table 2 (p. 13). The power of both tests declines with increases in k and decreases in the MA parameter. For samples equal to 50, the KPS test

12. It has been suggested to the authors that we could reduce the number of permutations by using the automatic data-dependent bandwidth procedure suggested by Andrews (1991). However, the approach used above is more commonly used by the applied researcher and is also that suggested by the tests' original authors. Furthermore, using the Andrews method may be problematic in the context of the KPS and Park tests since the regression residuals will have much more persistence than those in a DF or PP test, and Andrews notes that his estimator does poorly in such cases. Andrews and Monahan (1992) suggest the use of a prewhitened version of the Andrews estimator may be helpful, but we find that normally innocuous but necessary restrictions on the size of the prewhitening parameters are frequently binding for unit and near-unit roots.
13. See Hakkio and Rush (1991) for a discussion.
14. This is presumably because *k* should increase with *T* for consistency. Results in the literature provide various conditions for the increase of the truncation parameter as $T \to \infty$ that is sufficient for consistency. For example, depending on the underlying assumptions, the growth rate of *k* required for consistency can range from $T^{1/4}$ to $T^{1/2}$. However, Newey and West (1987) assert that the specification of an appropriate asymptotic growth rate for *k* provides little guidance for the choice of the truncation parameter in finite samples. This makes valid comparisons across finite samples for a given *k* difficult.

maintains power better over the truncation parameters. For example, for a truncation parameter equal to 8, the power of the J1 test declines to less than 2 per cent for all DGPs examined, while the power of the KPS test does not fall below 30 per cent. However, for larger samples both tests appear to maintain a reasonable level of power over *k*. Table 3 (p. 14) provides size-adjusted power calculations for each experiment. Again, the KPS test performs better than the J1 test for smaller sample sizes (50 observations). However, for larger sample sizes the J1 test appears to dominate, albeit slightly, the KPS test in size-adjusted power. This result is not really surprising once we recall the evidence suggesting that the J1 test is marginally more conservative than the KPS test in terms of empirical size.

In sum, as is the case with conventional unit-root tests, the KPS and J1 tests appear to suffer from size distortion and loss of power for certain data generation processes, and neither test clearly dominates the other. Both tests also appear sensitive to the specification of the truncation parameter.

4.0 Joint Tests for Unit Roots

KPS suggest using tests of the null hypothesis of stationarity in conjunction with tests of the null hypothesis of a unit-root. Using two tests means we may observe one of four possible outcomes. If the stationarity test rejects the stationary null and the unit-root test accepts the unit-root null, we would logically conclude that unit roots are present. If the former test accepts the null of stationarity and the latter rejects the presence of unit roots, we would logically conclude that the data are stationary. If neither test rejects its null, we could conclude that the data are simply not sufficiently informative to distinguish between stationarity and a unit root. Finally, if both tests reject their nulls, we may suspect some kind of misspecification, since at least one of the tests must be suffering from Type I error. Note that the latter two situations do not allow us to draw conclusions about the presence of unit roots.

To determine the value of using a joint test as opposed to a single traditional unit-root test, one must weigh the likelihood of producing conflicting and therefore inconclusive results against the reduced probability of an incorrect conclusion. Since this in turn depends on the relative likelihood of encountering different DGPs and on the value one places on avoiding Type I or Type II errors, no definitive answer is possible. However, the results we present below should be enough to show applied researchers the kind of trade-offs involved. For compactness, we focus exclusively on results using the KPS test.

In Table 4 (p. 15) and Table 5 (p. 16) we show the results of using a joint testing strategy on the same range of DGPs considered in the previous section. We use the same Monte Carlo framework described above to generate test statistics for both KPS and PP tests for each repetition.[15] We then cross-tabulate these results for easier interpretation, comparing the results of a single-test strategy using just the PP test to the joint testing strategy. They show the frequency of incorrect conclusions using the PP test along side the frequency of incorrect, inconclusive and correct conclusions using the PP and KPS tests jointly.[16]

Table 4 considers unit-root DGPs. We see that for non-negative MA parameters, both the single and the joint testing procedures produce few incorrect conclusions of stationarity, with the joint test always being more conservative than the single test. However, for negative MA parameters, the joint procedure reduces (often greatly) the number of incorrect conclusions produced by the single test. For example, with $T = 100$, $k = 4$ and $\theta = -0.8$, the frequency of such errors falls from 99.05 per cent with the PP test to 28.36 per cent with the KPS-PP test. The actual reduction tends to increase with T and decrease with k, which accords with the behaviour of the KPS test reported in Table 2 (p. 13). This comes at the cost of the ability to conclude that unit roots are indeed present in some cases when the MA parameter is non-negative. We see that while the frequency of the unit-root conclusion exceeds 99 per cent in several cases, it drops to just under 50 per cent for small T and large k. This inability to correctly detect the presence of unit roots is due to the KPS test's lack of power, as we can see by comparing the frequencies in the Unit Root

15. We chose the PP test over the DF test, since the PP and KPS approaches use the same long-run variance estimators. Further, we use the normalized bias version of the PP test statistic, as Campbell and Perron (1991) assert that it is more powerful than the standard regression "t-test". The PP Monte Carlo experiment results are available from the authors upon request.
16. Henceforth, when the PP test is used in conjunction with the KPS test, we refer to this as the KPS-PP test. A more complete breakdown of test results is available from the authors.

columns in Table 4 (p. 15) with the power figures for the KPS test in Table 2.

Table 5 (p. 16) shows the results for the stationary DGPs. Again, we find that the results differ depending on whether we consider negative or non-negative values of the MA parameter. For negative values, the single and the joint testing procedure produce nearly identical results. Even in 10,000 trials, many of these experiments failed to produce a single case of a conclusion that unit roots were present, and the frequency of such false conclusions is always below 3 per cent. For non-negative MA parameters, however, the single and joint tests give very similar numbers of incorrect conclusions in larger samples ($T = 200$), while for small samples and large truncation lengths ($k = 4$ and 8) the joint test produces far fewer incorrect conclusions than the single test. For example, with $T = 50$, $\theta = 0.5$ and $k = 4$, the single test fails to reject the unit-root null in 91.69 per cent of all trials, while the joint test admits a unit-root result in only 28.94 per cent of the cases. Again, this gain is offset by the inability to conclude that the DGP is stationary for a significant part of the parameter space, namely, in larger samples with negative MA parameters. While the PP test rejects the presence of a unit root over 97 per cent of the time in all such cases, the frequency of the stationary conclusion for the KPS-PP test varies from over 90 per cent to under 25 per cent, the frequency falling as T increases and θ and k decrease. Again, the inability to correctly detect the stationarity of the DGP is due to the KPS test's severe size distortion, as we can see by comparing the results in the Stationary Column in Table 5 with the size figures for the KPS test in Table 1 (p. 12).

As we stated at the outset of this section, whether the single or joint testing approach is superior in any well-defined sense will depend on the specific application and the relative importance the researcher places on Type I and Type II errors. However, an illustrative way to summarize our results would be to use a Bayesian approach and consider the *ex ante* and *ex post* probabilities of a unit root's presence. For convenience, we consider the simple case of ARMA(1,1) DGPs, and our priors place a probability of 0.5 on $\rho = 0.85, 1.0$ and 0.2 on $\theta = 0.8, 0.5, 0.0, -0.5, -0.8$. In other words, we examine only the DGPs considered in our Monte Carlo experiments, with each one thought to be equally likely. Table 6 (p. 17) shows the resulting *ex post* probabilities of having a true unit-root DGP conditional on the tests indicating the presence of a unit root (or in the case of the PP test, simply failing to reject the null). Table 7 (p. 17) shows the corresponding probabilities of having a truly stationary DGP conditional on the tests indicating stationarity.

Table 6 shows that in large samples ($T = 200$) there is little difference between the PP and KPS-PP tests when detecting a unit-root. Both the single and joint tests allow us to have similarly high confidence in our results, particularly when we have $k > 0$. In short data samples ($T = 50$), the single test does not make us any more confident about our results, moving our *ex post* probabilities little from our prior of 0.5, while the joint test allows us to be more confident of our results, particularly for large k. Table 7 shows significant differences between the single and joint testing procedure in all sample lengths when we conclude that the series is stationary. Based on the single test, we would believe this conclusion to be true between 60 and 75 per cent of the time for most choices of k and T, while with the joint test these probabilities rise to the 70 to 95 per cent range. These results suggest that the joint testing procedure allows us to be most confident about our results when it indicates our series are stationary, or when it indicates our series have a unit root for small samples and large k. However, we should expect the joint test to produce frequently inconclusive results. As is shown in Table 8 (p. 17), using the same priors used to construct Table 6 and Table 7, we should expect such results roughly 25 to 50 per cent of the time, but particularly in short samples with large k and in large samples with small k.

Some may suspect that the superior reliability of the joint test is simply the result of comparing it to the PP test, which we know suffers from severe size distortion in the presence of negative MA parameters.[17] Hence, we repeat the single and joint testing procedures comparison by replacing the PP test with the augmented Dickey-Fuller (ADF) test.[18] The results (Tables 9 and 10, p. 18) again suggest that the joint testing procedure will reduce the probability of an erroneous conclusion, especially for smaller samples and stationary DGPs.

---

17. See Schwert (1987,1989).
18. In the following experiment, we examine only the case for an arbitrary selected lag length of four.

We also examine whether the reduced probability of incorrect conclusions is a result of using two tests with opposite null hypothesis or simply a result of using two tests. To this end we compare the performance of the KPS-ADF test to the PP-ADF test (see Tables 9 and 10). For unit-root processes there is little disagreement between the ADF and PP tests except for large negative MA cases. Here the KPS-ADF test gives fewer stationarity conclusions than the PP-ADF test. However, for stationary processes the KPS-ADF test leads to appreciably fewer incorrect conclusions of unit roots, especially for smaller samples. For example, with $\rho = 0.85$, $\theta = 0.0$ and $T = 50$, the PP-ADF test will result in a unit-root conclusion about 69.3 per cent of the time, whereas the KPS-ADF test reduces this to 27.7 per cent.

5.0 Concluding Remarks

Both the KPS and J1 tests appear to suffer from size distortion and loss of power for certain data generating processes, and neither test clearly dominates the other. Both tests also appear sensitive to the specification of the truncation parameter. Examining the consequences of using the KPS test in conjunction with standard unit-root tests suggests that the frequency of incorrect conclusions can often be reduced relative to that of a single standard unit-root test. One way to summarize these results is in a simple Bayesian framework. We find that the joint testing approach gives the most reliable results when the joint test indicates that the data are stationary or that the data have a unit root for small samples and large k.

Table 1:
Empirical Size Calculations
Kwiatkowski, Phillips and Schmidt (KPS) and Park (J1) Tests
DGP: $y_t = 0.85y_{t-1} + \vartheta_t + \theta\vartheta_{t-1}$

k		0		1		2		4		8	
θ		KPS	J1	KPS	J1	KPS	J1	KPS	J1	KPS	J1
-0.8	T=50	12.29	11.87	10.16	9.25	8.87	6.44	7.53	2.99	3.86	1.34
	T=100	13.90	12.81	12.63	10.77	11.42	9.84	9.74	7.33	7.05	4.52
	T=200	15.44	14.42	13.74	12.71	13.00	11.16	11.28	8.82	8.53	6.87
-0.5	T=50	59.22	67.63	44.96	54.30	34.71	40.83	23.38	17.80	10.53	1.24
	T=100	68.96	65.48	52.77	50.40	41.96	39.34	27.87	23.50	16.71	7.28
	T=200	75.44	64.57	58.04	49.02	45.58	38.01	32.20	25.11	18.99	11.58
0.0	T=50	81.57	88.36	58.95	72.34	44.52	56.66	28.86	26.08	12.74	1.36
	T=100	88.25	81.24	67.14	61.40	52.30	48.03	33.04	27.28	18.84	7.73
	T=200	91.36	78.11	70.14	58.19	55.05	45.55	36.58	28.54	20.43	13.09
0.5	T=50	85.94	89.80	61.28	74.29	45.82	60.19	29.27	27.11	12.43	1.01
	T=100	90.36	83.15	69.33	64.74	52.01	49.56	34.43	27.76	19.09	8.71
	T=200	93.03	78.98	73.12	60.39	57.20	46.62	36.38	29.19	20.50	13.35
0.8	T=50	85.70	90.66	62.42	75.30	45.91	59.27	29.76	27.14	12.78	1.34
	T=100	90.96	83.50	69.47	64.53	52.55	49.45	33.41	28.12	19.35	8.65
	T=200	93.68	79.47	73.00	60.38	56.19	46.28	37.94	29.66	21.18	12.84

Note: The values in Tables 1 through 10 are in percentage form.

Table 2:
Empirical Power Calculations
Kwiatkowski, Phillips and Schmidt (KPS) and Park (J1) Tests
DGP: $y_t = y_{t-1} + \vartheta_t + \theta\vartheta_{t-1}$

k		0		1		2		4		8	
θ		KPS	J1	KPS	J1	KPS	J1	KPS	J1	KPS	J1
-0.8	T=50	61.63	63.74	56.63	57.20	52.66	51.14	44.69	32.75	31.18	1.51
	T=100	85.81	85.44	82.73	81.88	79.10	78.73	71.64	71.20	60.90	51.74
	T=200	97.42	95.33	95.97	95.54	94.03	91.97	90.44	88.48	80.64	80.79
-0.5	T=50	89.64	93.76	81.05	89.29	73.51	83.76	62.85	62.79	44.01	1.68
	T=100	98.40	96.24	94.76	93.65	90.95	89.54	80.45	81.24	67.31	61.15
	T=200	99.81	98.41	99.23	96.57	98.17	95.20	93.99	91.39	85.19	83.79
0.0	T=50	96.13	97.85	86.93	93.75	78.04	88.72	65.15	70.82	45.66	1.92
	T=100	99.40	97.74	96.19	94.60	91.75	90.77	82.17	82.41	67.97	62.61
	T=200	99.95	98.50	99.52	97.32	98.48	94.64	94.90	92.18	84.75	84.46
0.5	T=50	97.16	98.23	87.79	94.49	77.53	89.26	65.24	71.33	46.43	1.96
	T=100	99.63	98.14	96.85	94.45	92.47	91.37	82.25	83.07	68.05	63.73
	T=200	99.99	98.76	99.55	96.88	98.39	95.18	94.67	91.73	85.28	84.46
0.8	T=50	96.84	98.36	87.87	94.54	78.40	89.44	66.40	71.96	46.49	1.90
	T=100	99.49	98.03	96.34	94.62	92.04	90.56	82.48	83.39	68.31	62.90
	T=200	99.96	98.58	99.55	97.16	98.76	95.16	94.81	91.80	85.26	84.25

Table 3:
Size-adjusted Power Calculations
Kwiatkowski, Phillips and Schmidt (KPS) and Park (J1) Tests
DGP: $y_t = y_{t-1} + \vartheta_t + \theta\vartheta_{t-1}$

k		0		1		2		4		8	
θ		KPS	J1	KPS	J1	KPS	J1	KPS	J1	KPS	J1
-0.8	T=50	49.46	52.30	47.74	48.84	44.68	48.19	40.11	39.95	34.28	10.21
	T=100	76.94	79.43	73.21	76.13	69.53	73.48	64.58	67.23	57.34	53.10
	T=200	92.26	93.06	91.43	91.06	88.32	89.07	84.50	86.35	74.94	78.69
-0.5	T=50	47.97	55.35	44.13	52.93	42.58	49.10	39.19	40.76	32.92	10.15
	T=100	67.56	75.82	62.68	72.04	59.28	69.28	55.68	64.78	50.99	55.77
	T=200	85.64	88.57	79.85	85.85	77.53	84.96	73.95	82.40	69.41	77.82
0.0	T=50	42.99	51.49	40.02	47.33	39.24	45.68	38.24	38.75	30.84	9.76
	T=100	59.71	68.53	58.25	68.03	56.50	65.27	55.39	63.06	51.20	56.75
	T=200	78.58	83.73	74.59	83.52	74.87	81.48	71.00	80.59	67.46	77.82
0.5	T=50	40.19	47.20	40.09	48.00	38.14	45.35	35.98	37.10	32.63	11.16
	T=100	59.25	67.44	58.22	67.28	57.15	65.74	54.15	63.44	50.22	56.75
	T=200	75.41	81.80	73.38	82.20	74.70	81.99	71.49	80.06	68.29	77.19
0.8	T=50	39.03	48.28	39.09	47.49	39.89	44.13	37.03	39.01	31.70	10.28
	T=100	59.44	67.07	56.99	67.52	56.97	64.93	54.64	63.11	49.45	56.02
	T=200	74.05	82.59	74.26	81.88	73.25	81.68	70.72	80.41	68.32	77.17

Table 4:
Comparison of Conclusions Given by the PP Test and the KPS and PP Tests
DGP: $y_t = y_{t-1} + \vartheta_t + \theta\vartheta_{t-1}$

		PP Stationary			PP-KPS Stationary			PP-KPS Inconclusive			PP-KPS Unit Root		
θ	k	T= 50	T=100	T=200	T= 50	T=100	T=200	T= 50	T=100	T=200	T= 50	T=100	T=200
-0.8	0	99.55	99.86	99.97	38.37	14.19	2.58	61.18	85.67	97.39	0.45	0.14	0.03
	1	98.61	99.22	99.09	43.37	17.27	4.03	55.24	81.95	95.06	1.39	0.78	0.91
	2	98.57	98.83	98.67	47.34	20.90	5.97	51.23	77.93	92.70	1.43	1.17	1.33
	4	98.85	99.05	98.46	55.31	28.36	9.56	43.54	70.69	88.90	1.15	0.95	1.54
	8	99.66	99.64	99.12	68.82	39.10	19.36	30.84	60.54	79.76	0.34	0.36	0.88
-0.5	0	57.66	63.55	65.96	10.30	1.60	0.19	47.42	61.95	65.77	42.28	36.45	34.04
	1	47.76	49.17	47.55	17.50	5.21	0.77	31.71	43.99	46.78	50.79	50.80	52.45
	2	48.13	46.53	42.06	22.49	8.67	1.83	29.64	38.24	40.23	47.87	53.09	57.94
	4	52.59	49.60	42.46	30.94	17.50	5.79	27.86	34.15	36.89	41.20	48.35	57.32
	8	59.05	56.00	47.90	44.75	28.62	13.69	25.54	31.45	35.33	29.71	39.93	50.98
0.0	0	4.61	4.69	5.23	1.66	0.44	0.05	5.16	4.41	5.18	93.18	95.15	94.77
	1	5.01	5.11	5.06	3.33	1.68	0.38	11.42	5.56	4.78	85.25	92.76	94.84
	2	4.83	5.63	4.69	3.74	2.77	0.81	19.31	8.34	4.59	76.95	88.89	94.60
	4	5.73	5.92	5.35	4.99	4.00	1.87	30.60	15.75	6.71	64.41	80.25	91.42
	8	4.06	6.12	5.96	3.91	5.25	3.82	50.58	27.65	13.57	45.51	67.10	82.61
0.5	0	0.23	0.35	0.40	0.11	0.05	0.00	2.85	0.02	0.41	97.04	99.93	99.59
	1	1.67	1.60	1.62	1.22	0.63	0.18	11.44	3.49	1.71	87.34	95.88	98.11
	2	1.87	2.21	2.40	1.54	1.30	0.62	21.26	7.14	2.77	77.20	91.56	96.61
	4	1.31	2.49	2.99	1.29	1.95	1.46	33.44	16.34	5.40	65.22	81.71	93.14
	8	0.49	1.63	2.93	0.48	1.56	2.08	53.10	30.46	13.49	46.42	67.98	84.43
0.8	0	0.21	0.25	0.25	0.16	0.07	0.02	3.05	0.62	0.25	96.79	99.31	99.73
	1	1.20	1.32	1.78	0.91	0.54	0.16	13.51	3.90	1.91	85.58	95.56	97.93
	2	1.51	2.16	1.98	1.33	1.31	0.46	20.45	7.50	2.30	78.22	91.19	97.24
	4	1.09	2.12	2.75	1.03	1.56	1.25	32.63	16.52	5.44	66.34	81.92	93.31
	8	0.33	1.30	2.50	0.32	1.26	1.83	53.20	30.47	13.58	46.48	68.27	84.59

Table 5:
Comparison of Conclusions Given by the PP Test and the KPS and PP Tests
DGP: $y_t = 0.85y_{t-1} + \vartheta_t + \theta\vartheta_{t-1}$

		PP Unit Root			PP-KPS Unit Root			PP-KPS Inconclusive			PP-KPS Stationary		
θ	k	T= 50	T=100	T=200	T= 50	T=100	T=200	T= 50	T=100	T=200	T= 50	T=100	T=200
-0.8	0	0.00	0.00	0.00	0.00	0.00	0.00	12.29	13.90	15.44	87.71	86.10	84.56
	1	0.00	0.00	0.00	0.00	0.00	0.00	10.16	12.63	13.74	89.84	87.37	86.26
	2	0.00	0.00	0.00	0.00	0.00	0.00	8.87	11.42	13.00	91.13	88.58	87.00
	4	0.00	0.00	0.00	0.00	0.00	0.00	7.53	9.74	11.28	92.47	90.26	88.72
	8	0.00	0.00	0.00	0.00	0.00	0.00	3.86	7.05	8.53	96.14	92.95	91.47
-0.5	0	1.19	0.00	0.00	1.17	0.00	0.00	58.07	68.96	75.44	40.76	31.04	24.56
	1	2.93	0.00	0.00	2.55	0.00	0.00	42.79	52.77	58.04	54.66	47.23	41.96
	2	2.69	0.00	0.00	2.19	0.00	0.00	33.02	41.96	45.58	64.79	58.04	54.42
	4	1.79	0.00	0.00	1.19	0.00	0.00	22.79	27.87	32.20	76.02	72.13	67.80
	8	0.99	0.00	0.00	0.43	0.00	0.00	10.66	16.71	18.99	88.91	83.29	81.01
0.0	0	72.20	22.28	0.02	64.71	21.87	0.02	24.35	66.79	91.34	10.94	11.34	8.64
	1	69.05	22.73	0.03	48.10	20.05	0.03	31.80	49.77	70.11	20.10	30.18	29.86
	2	67.32	21.48	0.01	37.33	16.05	0.07	37.18	41.68	55.01	25.49	42.27	44.92
	4	67.80	20.74	0.09	25.33	11.27	0.07	46.00	31.24	36.53	28.67	57.49	63.40
	8	73.84	19.37	0.12	11.24	7.03	0.09	64.10	24.15	20.37	24.66	68.82	79.54
0.5	0	96.91	81.40	14.93	84.46	76.26	14.76	13.93	19.24	78.44	1.61	4.50	6.80
	1	88.73	52.43	1.89	58.41	42.78	1.80	33.19	36.20	71.41	8.40	21.02	26.79
	2	87.75	47.49	1.11	44.29	31.62	0.98	44.99	36.26	56.35	10.72	32.12	42.67
	4	91.69	49.73	1.61	28.94	23.89	1.19	63.08	36.38	35.61	7.98	39.73	63.20
	8	97.85	63.43	2.75	12.37	16.87	1.44	85.54	48.78	20.37	2.09	34.35	78.19
0.8	0	98.06	87.78	25.11	84.99	81.97	24.83	13.78	14.80	69.13	1.23	3.23	6.04
	1	90.83	57.49	2.23	60.29	46.99	2.15	32.67	32.98	70.93	7.04	20.03	26.92
	2	89.97	51.49	1.55	44.63	34.32	1.28	55.37	35.40	55.18	8.75	30.28	43.54
	4	93.58	52.86	1.91	29.54	24.31	1.47	64.26	37.65	36.91	6.20	38.04	61.62
	8	98.80	70.61	3.80	12.78	17.66	2.12	86.02	54.64	20.74	1.20	27.70	77.14

Table 6:
Bayesian *Ex Post* Probabilities of a Unit Root
Conditional on Significant Evidence of a Unit Root

	Phillips-Perron Test			PP-KPS Test		
k	T=50	T=100	T=200	T=50	T=100	T=200
0	55.72	63.38	89.12	58.353	64.76	89.23
1	57.89	72.15	98.81	64.70	75.35	98.86
2	58.21	74.10	99.24	68.68	79.90	99.33
4	57.19	73.43	98.97	73.71	83.14	99.20
8	55.34	68.61	98.08	82.06	85.43	98.82

Table 7:
Bayesian *Ex Post* Probabilities of a Stationary Series
Conditional on Significant Evidence of Stationarity

	Phillips-Perron Test			PP-KPS Test		
k	T=50	T=100	T=200	T=50	T=100	T=200
0	58.81	64.65	72.80	73.76	89.28	97.87
1	61.70	70.16	76.17	73.08	89.04	97.46
2	61.96	70.96	76.85	72.44	87.79	96.57
4	60.57	70.29	76.56	69.31	84.80	94.53
8	58.28	67.79	75.69	64.30	80.21	90.90

Table 8:
Probabilities of an Inconclusive Result with the PP-KPS Test

k	T=50	T=100	T=200
0	24.21	33.64	49.88
1	27.39	32.32	43.45
2	32.13	30.59	36.77
4	37.17	29.63	29.59
8	46.34	33.19	24.47

Table 9:
Comparison of Conclusions Given by the ADF Test and the KPS and ADF Tests
DGP: $y_t = y_{t-1} + \vartheta_t + \theta\vartheta_{t-1}$

	ADF			PP-ADF			KPS-ADF			KPS-ADF		
	Stationary			Stationary			Stationary			Unit Root		
θ	T=50	T=100	T=200	T=50	T=100	T=200	T=50	T=100	T=200	T=50	T=100	T=200
-0.8	15.89	28.16	36.03	15.88	28.16	36.03	14.31	17.37	8.45	43.37	59.66	62.79
-0.5	4.31	5.54	5.14	3.71	5.06	4.86	3.23	2.92	1.64	61.98	78.39	91.13
0.0	4.69	5.17	4.57	1.33	2.41	2.49	3.12	2.52	1.39	63.82	80.03	91.59
0.5	4.05	4.13	4.70	0.60	1.12	2.06	2.80	2.16	1.18	64.22	80.87	91.20
0.8	3.58	3.46	3.68	0.36	0.86	1.46	2.14	1.51	0.98	63.79	80.57	91.98

Table 10:
Comparison of Conclusions Given by the ADF Test and the KPS and ADF Tests
DGP: $y_t = 0.85y_{t-1} + \vartheta_t + \theta\vartheta_{t-1}$

	ADF			PP-ADF			KPS-ADF			KPS-ADF		
	Unit Root			Unit Root			Unit Root			Stationary		
θ	T=50	T=100	T=200	T=50	T=100	T=200	T=50	T=100	T=200	T=50	T=100	T=200
-0.8	54.28	3.20	0.00	0.00	0.00	0.00	6.28	1.08	0.00	44.94	88.27	88.85
-0.5	85.80	49.81	2.29	1.79	0.00	0.00	22.67	19.67	1.44	13.46	41.94	68.11
0.0	89.00	64.08	8.97	63.93	19.48	0.04	27.72	27.17	5.35	10.27	30.17	61.14
0.5	90.49	69.37	13.09	85.64	46.34	1.15	28.93	29.96	7.84	8.95	25.80	57.98
0.8	93.29	77.52	21.59	89.24	50.56	1.41	28.44	30.56	12.08	6.29	19.51	54.50

References

Abuaf, N. and P. Jorion. 1990. "Purchasing power parity in the long run." *Journal of Finance* 45: 157-74.

Andrews, D. W. K. 1991. "Heteroskedasticity and autocorrelation consistent covariance matrix estimation," *Econometrica* 59: 817-58.

Andrews, D. W. K. and J. C. Monahan. 1992. "An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator." *Econometrica* 60: 953-66.

Campbell, J. Y. and P. Perron. 1991. "Pitfalls and opportunities: What macroeconomists should know about unit roots." In *National Bureau of Economic Research Macroeconomic Annual 1991*, edited by O. J. Blanchard and S. Fischer. (Cambridge: MIT Press).

DeJong, D. N., J. C. Nankervis, N. E. Savin and C. H. Whiteman. 1992. "Integration versus trend stationarity in time series." *Econometrica* 60: 423-33.

Diba, B. T. and H. I. Grossman. 1988. "Explosive rational bubbles in stock prices?" *American Economic Review* 78: 520-30.

Dickey, D. A. and W. A. Fuller. 1979. "Distribution of the estimators for autoregressive time series with a unit root." *Journal of the American Statistical Association* 74: 427-31.

Hakkio, C. S. and M. Rush. 1991. "Cointegration: How short is the long run?" *Journal of International Money and Finance* 10: 571-81.

Khan, J. A. and M. Ogaki. 1992. "A consistent test for the null of stationarity against the alternative of a unit root." *Economics Letters* 39: 7-11.

Kwiatkowski, D., P. C. B. Phillips and P. Schmidt. 1991. "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?" Discussion Paper No. 979. Cowles Foundation for Research in Economics at Yale University, (New Haven).

Nelson, C. R. and C. I. Plosser. 1982. "Trends versus random walks in macroeconomic time series: Some evidence and implications." *Journal of Monetary Economics* 10: 139-62.

Newey, W. K. and K. D. West. 1987. "A simple, positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix." *Econometrica* 55: 703-8.

Park, J. Y. 1990. "Testing for unit roots and cointegration by variable addition." In *Co-integration, Spurious Regressions, and Unit Roots: Advances in Econometrics*, edited by T. B. Fomby and G. F. Rhodes, Jr. (London: JAI Press).

Park, J. Y., S. Ouliaris and B. Choi. 1988. "Spurious regressions and tests for cointegration." Working Paper No. 8807. Cornell University Center for Analytic Economics, (Ithaca).

Phillips, P. C. B. 1991a. "To criticize the critics: An objective Bayesian analysis of stochastic trends." Discussion Paper No. 950. Cowles Foundation for Research in Economics at Yale University, (New Haven).

Phillips, P. C. B. 1991b. "Bayesian routes and unit roots: De rebus prioribus semper est disputandum." Discussion Paper No. 986. Cowles Foundation for Research in Economics at Yale University, (New Haven).

Phillips, P. C. B. and P. Perron. 1988. "Testing for a unit root in time series regressions." *Biometrika* 75: 335-46.

Rose, A. K. 1988. "Is the Interest Rate Stable?" *Journal of Finance* 43: 1095-1112.

Schwert, G. W. 1987. "Effect of model specification on tests for unit roots in macroeconomic data." *Journal of Monetary Economics* 20: 73-103.

Schwert, G. W. 1989. "Tests for unit roots: A Monte Carlo investigation." *Journal of Business & Economic Statistics* 7: 147-59.

Sims, C. A. 1988. "Bayesian skepticism of unit-root econometrics." *Journal of Economic Dynamics and Control* 12: 463-74.

Bank of Canada Working Papers

1991

91-1	The Productivity-Inflation Nexus Revisited: Canada, 1969-1988	F. Novin
91-2	Terms of Trade Movements in Major Industrial Countries, 1969-1989	S. Roger
91-3	Implementation of Monetary Policy with Same-Day Settlement: Issues and alternatives	D.Longworth and P. Muller
91-4	The Relationship Between Money, Output and Prices	F. Caramazza and C. Slawner
91-5	A Flexible, Forward-Looking Measure of Wealth	R. T. Macklem
91-6	The Predictability of Stock Market Regime: Evidence from the Toronto Stock Exchange	S. van Norden and H. Schaller

1992

92-1	Should the Change in the Gap Apppear in the Phillips Curve? Some Consequences of Mismeasuring Potential Output	D. Laxton, K. Shoom and R. Tetlow
92-2	Determinants of the Prime Rate: 1975-1989	S. Hendry
92-3	Is Hysteresis a Characteristic of the Canadian Labour Market? A Tale of Two Studies	S. Poloz and G. Wilkinson
92-4	Les taux à terme administrés des banques	J.-P. Caron
92-5	An Introduction to Multilateral Foreign Exchange Netting	W. Engert
92-6	Inflation and Macroeconomic Performance: Some Cross-Country Evidence	B. Cozier and J. Selody
92-7	Unit Root Tests and the Burden of Proof	R. Amano and S. van Norden

Single copies of Bank of Canada papers may be obtained from:

Publications Distribution
Bank of Canada
234 Wellington Street
Ottawa, Ontario, K1A 0G9