TESTING FOR A RANDOM WALK: A SIMULATION EXPERIMENT OF POWER WHEN THE SAMPLING INTERVAL IS VARIED

Pierre Perron*
Princeton University and C.R.D.E.

Econometric Research Program
Research Memorandum No. 336

Revised September 1988

*Assistant Professor, Department of Economics, Princeton University, Princeton, New Jersey 08544, and Research Associate, Centre de Recherche et Développement en Economique (Université de Montréal).

The material in this paper is drawn from Chapter 1 of the author's Ph.D. dissertation, Perron (1986). I wish to thank Peter C. B. Phillips and Robert J. Shiller for their useful comments. Funding is acknowledged from SSHRC (Canada) and FCAR (Québec).

Econometric Research Program
Princeton University
207 Dickinson Hall
Princeton, NJ 08544, USA

ABSTRACT

This paper analyses the power of various tests for the random walk hypothesis against AR(1) alternatives when the sampling interval is allowed to vary. The null and alternative hypotheses are set in terms of the parameters of a continuous time model. The discrete time representations are derived and it is shown how they depend on the sampling interval. The power is simulated for a grid of values of the number of observations and the span of the data available (hence for various sampling intervals). Various test statistics are considered among the following classes: a) test for a unit root on the original series and b) tests for randomness in the differenced series. Among class (b), we consider both parametric and nonparametric tests, the latter including tests based on the rank of the first-differenced series. The paper therefore not only provides information as to the relative power of these tests but also about their properties when the sampling interval varies. This work is an extension of Perron (1987) and Shiller and Perron (1985).

Key Words: Monte Carlo experiment, hypothesis testing, continuous time processes, test consistency.

1. INTRODUCTION

An interesting feature in time series analysis is that a given stochastic process can be parameterized in different ways corresponding to different specifications of the sampling frequency. Indeed, an infinity of combinations is theoretically possible by fixing different values of the number of observations T, the span of the data available S and the sampling frequency h. By definition, these parameters are related as T = S/h.

Consider, for example, the following simple continuous time Ornstein-Uhlenbeck process:

$$dy_t = -\gamma y_t dt + \sigma dw_t$$

where w_t is the unit Weiner process. The discrete time representation of the stochastic variable y_t is given by (see section 2):

$$y_{ht} = \beta_h y_{ht-1} + v_{ht}$$

where $\beta_h = \exp(-\gamma h)$ and h is the sampling interval. Suppose we wish to test the null hypothesis of a random walk, i.e. $\gamma = 0$ and hence $\beta_h = 1$ for all h. Under the null hypothesis, the discrete time autoregressive parameter β_h is independent of h. However, under the alternative hypothesis that the process is stationary ($\gamma > 0$), the autoregressive parameter β_h depends on the specified sampling interval h for any fixed value of the continuous time parameter γ.

The usual analysis concerning the power properties of various tests of the random walk hypothesis does not explicitly take into account the dependence of the power on the sampling interval. The autoregressive parameter is usually treated as fixed and the power analysed for various sample sizes. Indeed, the usual consistency criterion considers a test consistent if its power function converges to one as the sample size converges to infinity for any given fixed alternative β_h.

Yet, some interesting issues cannot be raised within this framework. For instance, does an increase in the sample size lead to the same increase in power if it is achieved by keeping a fixed span (i.e., reducing the sampling interval) or by increasing the span one for one with the number of observations (i.e. keeping the sampling interval fixed)? Is it possible to have higher power with fewer observations if these observations are spread out over a longer period? Is it possible that the power of some test statistics can actually be increased by simply deleting some observations (i.e. keeping the same span but with a larger sampling interval)?

These questions and others can be answered by an analysis of the power function of the test which explicitly takes into account the dependence of the autoregressive parameter β_h on the sampling interval h and evaluates the power directly in terms of the fixed alternative value of the continuous time parameter γ.

Similarly, asymptotic analysis can be used to shed some light on the above issues. Indeed, one can consider the consistency property of a test statistic for a given fixed alternative $\gamma(\gamma > 0)$ as the sample size increases to infinity allowing any path for the sampling interval. For example, one could consider the consistency of a test statistic as the sample size converges to infinity with a given fixed span S; i.e. when the sampling interval converges to zero at rate T. If a test is inconsistent under this path for h, the sampling interval, one could expect low power, in finite samples, with a data set sampled frequently. On the other hand, if a test is consistent allowing h to increase as the sample size increases (i.e., allow the span S to increase faster than T), then one could expect large power in a data set that extends over a long horizon even with a relatively few number of obserations.

The aim of the paper is to analyse the power properties of various tests of the random walk hypothesis against stationary alternatives in a context which allows for different sampling intervals. Section 2 introduces the continuous time model and considers its properties in more detail. The test statistics are presented in Section 3. We include statistics that test for a unit root in the original data and statistics that test for randomness in the first-differenced series. In the latter class, we include parametric as well as nonparametric statistics (including some nonparametric rank statistics). Hence, a by-product of our study is also a comparison of power among a wide number of statistics from different classes (13 statistics are considered in all). The simulation experiment to assess the power in finite samples is presented in section 4 and the results are discussed in section 5.

Section 6 presents some theoretical results concerning the consistency properties of the tests analysed when we allow an arbitrary path for the sampling interval as the sample size increases. These results draw heavily from the study of Perron (1987) which used methods first introduced by Phillips (1987a, 1987b and 1988) in the context of continuous asymptotic records for integrated series and asymptotic analyses of near-integrated series. Section 7 presents some concluding comments and a mathematical appendix contains the proofs of the theorems in section 6.

Some of the results obtained are as follows. First, the power depends more importantly on the span of the data rather than the number of observations per se for all statistics considered. It is preferable to have a large span of data even, in most cases, if this entails a smaller number of observations available. Second, there is a notable difference between tests using the original level of the series and tests based on testing for randomness in first-differenced data. In the latter case, too many observations, for a given fixed span, may destroy the power. It is shown, in particular, that as the number of observations increases, keeping a fixed span, the power converges to the size of the test. It may be the case that higher power can be obtained by deleting observations while keeping the span fixed. This feature is not present for the class of tests based on the original undifferenced series. In that case, more observations always lead to higher power though the marginal contribution of each additional observation is quickly declining.

2. THE NULL AND ALTERNATIVE HYPOTHESES

The simplest and most sensible way to approach the problem of differential sampling interval and its effect on the properties of test statistics is to consider the general case of correlation in continuous time records. This is a limiting case but it can be argued that, with the increasing availability of frequent sampling of data in some areas of economics, it may represent an interesting approximation. We therefore posit a process which occurs continuously. Our null hypothesis is that the changes in the random variable of interest are independently and identically distributed and the alternative is that the process is mean reverting (i.e. the true process is one of correlation where the variable tends to return to its mean value). These null and alternative hypotheses are succinctly represented by the stochastic differential equations:

$$H_0: \quad dy(t) = \sigma dw(t) \qquad t > 0 \ , \ y(0) = 0 \tag{2.1}$$

$$H_1: \quad dy(t) = -\gamma y(t)dt + \sigma dw(t), \ -\infty < t < \infty \quad , \quad \gamma > 0 \tag{2.2}$$

where $w(t)$ is a unit Weiner process and γ and σ are constant. Under the null hypothesis $y(0)$ is fixed at 0 while, since the alternative hypothesis specify a stationary process, we can equivalently specify H_1 as holding for $t \geq 0$ with $y(0)$ having the stationary distribution (see below). The null hypothesis is, in fact, that $\gamma = 0$.

These systems of stochastic diffential equations can be solved in order to derive discrete time representations of the processes. Let h be the sampling interval. Consider first the alternative hypothesis where $y(t)$ is simply an Ornstein-Uhlenbeck process having a unique solution of the form (see e.g. Arnold, 1974, pp. 134-5):

$$y(th) = e^{-\gamma ht}y(0) + \sigma \int_0^{ht} e^{-\gamma(ht-s)}dw(s) \quad , \quad t \geq 0$$

Simple manipulations (see, for example, Arnold, p. 134) yield the following discrete time representation of the alternative hypothesis (imposing stationarity):

$$H_1: \quad y(ht) = e^{-\gamma h}y\big(h(t-1)\big) + v(ht) \qquad t = 1, 2, \ldots \tag{2.3}$$

where $y(0) \sim N(0,\sigma^2/2\gamma)$ and $v(ht) \sim N\left(0,\ \sigma^2(1-\exp(-2\gamma h))/2\gamma\right)$. Under the null hypothesis H_0, the unique solution to the stochastic differential equation (2.1) is:

$$y(ht) = \sigma \int_0^{ht} dw(s) = \sigma \int_0^{h(t-1)} dw(s) + \sigma \int_{h(t-1)}^{ht} dw(s).$$

Let $u(th) = \sigma \int_{h(t-1)}^{ht} dw(s)$. From the properties of the Brownian motion $w(t)$, it is easy to deduce that $\mathrm{Var}\left(u(th)\right) = h\sigma^2$ (see Arnold, 1974 and Bergström, 1984). Therefore, in discrete time, we have the following null hypothesis:

$$H_0: \quad y(ht) = y\left(h(t-1)\right) + u(ht) \qquad t = 1, 2, \ldots \qquad (2.4$$

where $y(0) = 0$ and $u(ht) \sim N(0,\ h\sigma^2)$. If we let $\beta_h = \exp(-\gamma h)$, the alternative hypothesis has the form

$$H_1: \quad y(ht) = \beta_h y\left(h(t-1)\right) + v(ht) ,$$

and a test of the null hypothesis that $\gamma = 0$ is equivalent in discrete time to a test that $\beta_h = 1$, i.e. the random walk hypothesis; the test being carried out against an alternative that the process is a stationary first-order autoregressive process.

Of course, in practice only a finite amount of data is available, say T. In the following, we will denote by S the span of the data available, where $S = hT$. Therefore, in discrete time, the index t is in the range $t = 0, 1, \ldots, T = S/h$.

Four of the tests we study deal directly with the sample series $\{y(ht)\}_{t=0}^{T}$ but the rest of them treat the series $\Delta y(ht) \equiv y(ht) - y\left(h(t-1)\right)(t=1,\ldots,T)$. Indeed under the null hypothesis that $\beta_h = 1$ (for all h) we have that $\Delta y(ht)$ is a sequence of independent and identically distributed random variables so that we can simply apply various tests of randomness readily available in the literature. It is therefore of interest to derive the null and alternative hypotheses in terms of the sample of first-differences $\{\Delta y(ht)\}_{t=1}^{T}$. In what follows, we will omit the suffix h and simply write Δy_t and y_t.

Of course, under the null hypothesis $\Delta y_t = u_t (t = 1, \ldots, T)$ where u_t is a sequence of iid normal random variables with mean 0 and variance $h\sigma^2$. Under the alternative hypothesis, the sequence of first differences $\{\Delta y_t\}$ is an ARMA(1,1) process with a moving average parameter on the unit circle, i.e.

$$\Delta y_t = \beta_h \Delta y_{t-1} + v_t - v_{t-1}.$$

Some algebra yields the following representation for the k^{th} order autocorrelation of the sequence $\{\Delta y_t\}$, ρ_k:

$$\rho_k = (\tfrac{1}{2})(\beta_h - 1)\beta_h^{k-1} = (\tfrac{1}{2})(\exp(-\gamma h) - 1)\ \exp(-\gamma h(k - 1)).$$

The autocorrelations are negative at all lags and have a maximum in absolute value when k = 1. Note the fact that $\rho_k \to 0$ for all k as $h \to 0$. As $h \to \infty$, we have $\rho_1 \to -\frac{1}{2}$ and $\rho_k \to 0$ for all $k \geq 2$. These observations will be useful when interpreting the behavior of the power of the tests.

The framework considered here may seem overly restrictive with the normality and independence assumptions for the error sequence and the fact that no constant nor deterministic time trend are included. The latter could be relaxed by using different statistics than the ones presented in the next sections. Much the same conclusions regarding the behavior of the power functions as the sampling frequency is changed would remain. The present study is simply illustrative of some phenomena that occur in a more general context. The normality assumption has been relaxed in Perron (1986) and the conclusions are basically the same. Our study, however, cannot generalize to the case where additional correlation is present in the errors since the test statistics based on the sequence $\{\Delta y_t\}$ are indeed constructed for testing the null hypothesis that Δy_t is uncorrelated.

3. DESCRIPTION OF THE STATISTICS

In this section, we describe various test statistics which can be used to test the random walk hypothesis. Thirteen test statistics are described. Four of these tests are based upon the original series of T + 1 observations $y_0, \ldots, y_T$; namely the ordinary least square estimator in a regression of y_t against y_{t-1} and its associated t-statistic, the locally best invariant and maximal invariant tests. The other tests are based upon the series of first differences $\Delta y_t = y_t - y_{t-1}$ $(t = 1, \ldots, T)$. (See also Lepage and Zeidan (1981) and Girard (1983) who analysed some of the statistics described here in a different context.) In this section, we pay special attention to the asymptotic distribution of these tests and to the determination of the appropriate critical values to be used in the simulation study. For simplicity of notation, we suppress the dependence of the coefficients and variables on the parameter h, the sampling interval.

A. Test statistics based upon the series $\{y_t\}_0^T$

1) Normalized OLS coefficient:

An obvious test statistic to consider is the ordinary least squares estimates of the autoregressive parameter β, $\hat{\beta}$, in a regression of y_t against y_{t-1} (this is also the maximum likelihood estimator of β conditional upon y_0):

$$\hat{\beta} = \Sigma_{t=1}^T y_t y_{t-1} / \Sigma_{t=1}^T y_{t-1}^2$$

It is easy to verify that $\hat{\beta}$ is invariant with respect to σ^2, the variance of the errors term. Therefore it can be used directly as a basis for testing $\beta = 1$. It can be shown (see Fuller, 1976, lemma 8.5.1) that $\hat{\beta} - 1 = 0_p(T^{-1})$ so that $T(\hat{\beta} - 1)$ has a non-degenerate asymptotic distribution. This distribution has been tabulated using Monte Carlo methods by Dickey (1976) and a table of values can be found in Fuller (1976), Table 8.5.1, for certain values of T which, however, do not fit our experiment. We obtained the required critical values by simulation using 10 000 replications (see the methodology section) for all values of T considered in this study.

2) t-statistic on $\hat{\beta}$:

Another natural statistic to consider is the t-statistic associated with the previous regression
$t_{\hat{\beta}} = (\hat{\beta} - 1)(\Sigma_{t=1}^{T} y_{t-1}^{2})^{\frac{1}{2}}/\hat{\sigma}$ where $\hat{\sigma}^2 = (T-1)^{-1} \Sigma_{t=1}^{T}(y_t - \hat{\beta}y_{t-1})^2$.

The t-statistic is a consistent test of $\beta = 1$, i.e $t_{\hat{\beta}} = 0_p(1)$ (see Fuller, 1976; Lemma 8.5.2). Percentage points of the distribution of $t_{\hat{\beta}}$ have been evaluated by Monte Carlo method by Dickey (1976) and the values are tabulated in Table 8.5.2 in Fuller (1976). Since this table does not include exact values for the values of T which are of interest to us, these were estimated by Monte Carlo method using 10,000 replications (see the methodology section).

3) Locally best invariant:

King (1981) derived the locally best invariant (LBI) test for the hypothesis that $\beta = 1$ against $\beta < 1$, i.e. a one-sided test [see also Dufour and King (1986)]. This test consists in rejecting the null hypothesis for low values of d, where d is defined as:

$$d = y_T^2/\Sigma_{t=1}^{T}(y_t - y_{t-1})^2.$$

Under the null hypothesis, we can write d in terms of the innovation sequence $\{u_t\}$ as $d = u'ss'u/u'u$ where $s = (1,\ldots,1)'_{(Tx1)}$, $y_T = s'u$ and $u' = \{u_1, \ldots, u_T\}$. The $T \times T$ matrix ss' is of rank 1 and therefore has T - - 1 zero eigenvalues and one non-zero eigenvalue taking the value T. Therefore, we can also write d as $d = T\, \xi_1^2/\xi'\xi$ where $\xi = (\xi_1, \ldots, \xi_T)'$ and $\xi_i \sim IN(0,1)$. Finally, we note that d/T is of the form $X_1/(X_1 + X_2)$ where X_1 and X_2 are independent X^2 variates and thus d/T has a beta type I distribution with parameters $\frac{1}{2}$, $(T-1)/2$. These critical values where used in the simulation studies, and where obtained using the subroutine MDBETI from the IMSL library.

4) Uniformly most powerful invariant:

Bhargava (1986) extended the work of Sargan and Bhargava (1983) to provide uniformly most powerful invariant tests of the random walk hypothesis. As with the previous test of King the resulting test statistics are different for stationary and explosive alternatives. We shall be concerned here with the former only. In the case studied here, the statistic reduces to

$$R = \Sigma_{t=1}^{T}(y_t - y_{t-1})^2 / \Sigma_{t=0}^{T}(y_t - Y)^2, \ \overline{Y} = (T + 1)^{-1}\Sigma_0^T y_t$$

It is a ratio of quadratic forms in normal random variables and therefore the critical values can be evaluated with the Imhof routine. Since the critical values tabulated by Bhargava are not for sample sizes used in this study, we derived them independently using the Imhof routine as programmed by Koerts and Abrahamse (1969). Note that Bhargava provides the appropriate eigenvalues to be used as weights in this routine.

B. Tests based on the series $\Delta y_t = y_t - y_{t-1} (t = 1,\dots, T)$

We now consider test statistics based upon the series of first differences, denoted $\Delta y_t \equiv y_t - y_{t-1}$ $(t = 1, \dots, T)$. As seen in the previous section, under the null hypothesis that the series $\{y_t\}_0^T$ is a random walk, the series $\{\Delta y_t\}_1^T$ is a random sequence, i.e. $\Delta y_1, \dots, \Delta y_T$ are mutually uncorrelated. We can therefore apply a variety of test statistics to test this null hypothesis of randomness.

B-1. Parametric tests

5) The von Neumann ratio:

Von Neumann (1942) suggested as a statistic the ratio of successive mean squared differences to the variance:

$$V = (T/T-1) \ \Sigma_{t=2}^{T}(\Delta y_t - \Delta y_{t-1})^2 / \Sigma_{t=1}^{T}(\Delta y_t - \overline{\Delta y})^2 \text{ where } \overline{\Delta y} = T^{-1}\Sigma_{t=1}^{T}\Delta y_t$$

Under the null hypothesis, Von Neumann (1942) showed that with normal errors, $E(V) = 2T/(T - 1)$ and $VAR(V) = 4/T$. It is also shown that $(V - E(V))/\sqrt{Var(V)}$ is asymptotically $N(0,1)$. Since V is a ratio of quadratic forms, its exact distribution can be computed using the Imhof routine. This was done by Dufour and Perron (1985) for a wide range of values of T and percentage points. For T up to 64, we take the critical values from this study which also showed that the normal approximation is indeed very good for T greater than 60. Therefore, for values of T greater than 64 we determine the critical values from the asymptotic normal distribution.

6) First-order correlation coefficient:

Consider the first-order serial correlation coefficient defined by:

$$r = (T/T-1)\Sigma_{t=2}^{T} \Delta y_t \Delta y_{t-1} / \Sigma_{t=1}^{T} \Delta y_t^2.$$

Under the null hypothesis of randomness in the series $\{\Delta y_t\}$ and if the Δy_t's have a common probability density function with finite moments, then Moran (1948) shows that $E(r) = 0$. However, to obtain an exact value for the variance, the common p.d.f. must be assumed normal, and $Var(r) = T(T - 1)^{-1}(T + 2)^{-1}$. Anderson (1942) shows that, under normality of the errors, $[r - E(r)]/\sqrt{Var(r)}$ is asymptotically $N(0, 1)$. In this study, we used the critical values derived from the asymptotic distribution.

B-2. Nonparametric tests

We now discuss a variety of nonparametric tests and their distribution under the null hypothesis. A comment about the determination of the critical values is in order. In this section, all the test considered have the property that $[J-E(J)]/\sqrt{Var(J)}$ tends asymptotically to a $N(0,1)$ variable as $T \to \infty$, under the null hypothesis. For most statistics, the asymptotic approximation is very good even for quite small values of T. Nevertheless, there may be a significant discrepancy for values of T as small as 8, and even 16, studied in this paper. The exact distribution have been tabuled in most cases for such small sample sizes. However, due to the discreteness of the exact distribution we cannot get a critical value for which a test of size 0,05 can be constructed. A possible resolution of this problem would be to use randomized test procedures which would create a test of exact size 0,05. Since the main concern of this paper is the behavior of the power function as h tends to 0 with T increasing, it is not worthwhile to carry such a procedure. We therefore use the asymptotic critical values for all sample sizes. The effect on the size of the tests can be evaluated and the estimated sizes are presented with the power results.

B-2a. Tests based on the level of the series $\{\Delta y_t\}$

7) Turning points test:

We define the series $\Delta y_1, \ldots, \Delta y_T$ as having a turning point at t if $\Delta y_{t-1} > \Delta y_t$ and $\Delta y_t < \Delta y_{t+1}$ or if $\Delta y_{t-1} < \Delta y_t$ and $\Delta y_t > \Delta y_{t+1}$. The statistic of interest, say D, is simply the number of such turning points present in the series. That is,

$$D = \Sigma_{t=2}^{T-1} Y_i \text{ where } Y_i = 1 \text{ if there is a turning point at } i \text{ and } Y_i = 0 \text{ otherwise.}$$

If $T > 4$, we have $E(D) = 2(T - 2)/3$ and $VAR(D) = (16T - 29)/90$ (see Kendall and Stuart, 1976). It can also be shown that this test is equivalent to a test based upon the number of runs (see Mood, 1940) a version of which was studied in Shiller and Perron (1985). The turning point test is more powerful against cyclical alternative than for trend alternative for which it has very low power (see Lepage and Zeidan, 1981). Also, Knoke (1977) showed that the asymptotic relative efficiency of D with respect to r, the correlation coefficient, is 0.19 for alternatives which are first-order autoregressive in the Δy's.

8) Wald-Wolfowitz statistic:

Wald and Wolfowitz (1943) proposed the following transformation of the serial correlation coefficient

$$R = \Sigma_{t=2}^{T} \Delta y_t \Delta y_{t-1} + \Delta y_1 \Delta y_T$$

which is designed especially to test against serial behavior. They show that

$$E(R) = (S_1^2 - S_2)/(T - 1)$$

$$Var(R) = (T - 1)^{-1}[S_2^2 - S_4] + [(T - 1)(T - 2)]^{-1}$$
$$[S_1^4 - 4S_1^2S_2 + 4S_1S_3 + S_2^2 - 2S_4] - (T - 1)^{-2}[(S_1^2 - S_2)^2]$$

where $S_k = \Sigma_{t=1}^T \Delta y_t^k$. Under some regularity conditions $[R-E(R)]/[VAR(R)]^{\frac{1}{2}}$ is asymptotically normal.

B-2b. Tests based on the ranks of $\{\Delta y_t\}_1^T$

9) Rank correlation coefficient:

If we apply the ordinary correlation coefficient to the ranks of the series Δy_t, we can define the following rank correlation coefficient, where $\{R_t\}$ is the series of ranks associated with the series $\{\Delta y_t\}$.

$$r_k = (T/T - 1)\ \Sigma_{t=2}^T (R_t - \bar{R})(R_{t-1} - \bar{R})/\Sigma_{t=1}^T (R_t - \bar{R})^2.$$

Since $\Sigma_{t=1}^T R_t = \Sigma_{t=1}^T t = T(T + 1)/2$ and $\Sigma_{t=1}^T R_t^2 = \Sigma_{t=1}^T t^2 = T(T + 1)(2T + 1)/6$, using r_k is equivalent to using the statistic

$$K = \Sigma_{t=2}^T (R_t - \bar{R})(R_{t-1} - \bar{R}) \quad \text{where } \bar{R} = (T + 1)/2.$$

Knoke (1977) studied the statistic K and its power against autocorrelated alternatives. Now it can be shown that $E(K) = -(T^2-1)/12$ and $Var(K) = (T + 1)T^2 (T - 3)(5T + 6)/720$. Knoke also showed that the asymptotic relative efficiency of K with respect to the ordinary correlation coefficient is 0.91 for first-order autoregressive alternatives.

10) Rank version of von Neumann ratio:

Bartel (1982) studied the rank version of the Von Neumann ratio given by

$$G' = \Sigma_{t=2}^T (R_t - R_{t-1})^2/\Sigma_{t=1}^T (R_t - \bar{R})^2$$

where $\bar{R} = T^{-1}\Sigma_{t=1}^T R_t = (T + 1)/2$. Since the denominator is equal to $T(T^2 - 1)/12$, using G' is equivalent to using:

$$G = \Sigma_{t=2}^T (R_t - R_{t-1})^2.$$

Under H_0, Bartel obtained $E(G) = T(T^2 - 1)/6$ and $Var(G) = T(T + 1)(T - 2)(5T^2 - 2T - 9)/180$. The asymptotic normal approximation is adequate for $N \geq 25$. Bartel also concluded that the power of G compared favorably against first-order correlated alternatives.

C. Tests based on the ranks of $\Delta y_t \Delta y_{t+1}$

Dufour (1981) introduced a family of linear rank test statistics to test the independence of a sequence of random variables under the assumption that the marginal probability density function of the series is symmetric with zero median but without requiring the assumption that the series is identically distributed. These tests are aimed against alternatives of serial dependence.

Dufour motivates the statistics by noting that under the symmetry and independence assumptions, we have $\text{med}(\Delta y_t \Delta y_{t+k}) = 0$ $(t = 1, \ldots, T-k)$ where med $(\cdot)$ refers to the median. If there is positive serial dependence (at lag k) $\text{med}(\Delta y_t \Delta y_{t+k}) > 0$ and with negative serial dependence $\text{med}(\Delta y_t \Delta y_{t+k}) < 0$. The following tests are therefore aimed at testing the null hypothesis that med $(\Delta y_t \Delta y_{t+k}) = 0$ against the alternative hypothesis that med $(\Delta y_t \Delta y_{t+k}) \neq 0$, with $t = 1, \ldots T-k$ and k an integer $(1 < k < T)$. The family of rank statistics is described as follows:

$$S_k = \Sigma_{t=1}^{T-k} \mu(Z_t)\ a_{T-k}(R_t^+)$$

where $Z_t = \Delta y_t \Delta y_{t+k}$, R_t^+ = rank of $|Z_t|$ and $\mu(Z_t) = 1$ if $Z_t \geqslant 0$ and 0 otherwise. It is shown that the mean and variance of S_k under H_0 are given by $E(S_k) = (\frac{1}{2})\Sigma_{t=1}^{T} a_T(t)$ and $Var(S_k) = (1/4)\Sigma_{t=1}^{T} a_T^2(t)$. Furthermore, the distribution of S_k is symmetric about $E(S_k)$ and approximately normal under some regularity conditions.

The choice of the function $a_T(t)$ defines the particular statistics. In this study, we will analyse three statistics defined by i) $a_T(t) = 1$; ii) $a_T(t) = t$ and iii) $a_T(t) = \phi^{-1}((1 + t(T + 1)^{-1})/2)$, where ϕ^{-1} is the inverse of the N(0,1) c.d.f.

ii) Sign test:

If we let $a_T(t) = 1$ for all t $(t = 1, \ldots T)$, we have:

$$S_k^{(1)} = \Sigma_{t=1}^{T-k} \mu(Z_t)\ .$$

In this case, $S_k^{(1)}$ is the number of non-negative values in the sequence $Z_1, \ldots Z_{T-k}$, i.e. the statistic of the sign test applied to $Z_1, \ldots Z_{T-k}$. Under H_0, the exact distribution of $S_k^{(1)}$ is Bin(T-k, 1/2) where Bin(•) is the binomial distribution.

If we let k = 1, $S_1^{(1)}$ is the number of times consecutive Δy_t's have the same sign. Thus $T - S_1^{(1)}$, is the total number of runs in the sequence $\mu(\Delta y_1), \ldots \mu(\Delta y_T)$. We approximate the distribution using the asymptotic normality property with $E(S_1^{(1)}) = (T-1)/2$ and $Var(S_1^{(1)}) = (T-1)/4$.

12) <u>Wilcoxon signed-rank test:</u>

If we let $a_T(t) = t$, we get $S_K^{(2)} = \Sigma_{t=1}^{T} \mu(Z_t) R_t^+$ which is the sum of the ranks of the non-negative Z_t's. This test statistic is associated with the Wilcoxon signed-rank test for symmetry about zero when applied to $Z_1, \ldots, Z_T$. The exact distribution of S_K under H_0 is the same as the null distribution of the Wilcoxon test statistic. In this study, we let k = 1 and it can be verified that in this case $E(S_1^{(2)}) = T(T - 1)/4$ and $Var(S_1^{(2)}) = T(T - 1)(2T - 1)/24$.

13) <u>Van der Waerden test statistic:</u>

If we let $a_T(t) = \phi^{-1}((1 + t(T + 1)^{-1})/2)$ where ϕ^{-1} is the inverse of the cumulative distribution function of a N(0, 1) random variable, then we get the analog of the Van der Waerden (1952) test statistic:

$$S_1^{(3)} = \Sigma_{t=1}^{T-1} \phi^{-1}((1 + R_t^+(T + 1)^{-1})/2)\ \mu(Z_t).$$

Under H_0:

$$E(S_1^{(3)}) = \Sigma_{t=1}^{T-1} \phi^{-1}((1 + R_t^+(T + 1)^{-1})/2)/2$$

and $$Var(S_1^{(3)}) = \Sigma_{t=1}^{T-1}\ \phi^{-1}((1 + R_t^+(T + 1)^{-1})/2)^2/4.$$

4. METHODOLOGY FOR THE SIMULATIONS

All simulations were carried out on a CDC Cyber 915 at the Université de Montréal. The N(0,1) random deviates were obtained from the subroutine GGNML of the International Mathematical and Statistical Library (IMSL). The critical values for $\hat{\beta}$ and $t_{\hat{\beta}}$ were obtained using 10 000 replications. The starting value y(0) was set at 0. From the ordered sequence (using the subroutine VSRTA) the 5.0 percentage point was taken. Hence, all tests considered here are one-tailed 5 % size tests against stationary alternatives. The power estimates were obtained using 2 000 replications. Hence the standard error of any entry in the Tables is $\sqrt{P(1 - P)/2\ 000}$ if P is the true power, which gives a maximum standard error of 0,0112 (when $P = \frac{1}{2}$).

Since the tests statistics are invariant to the value of σ^2, we let $\sigma^2 = 1/h$ under the null hypothesis and $\sigma^2 = 2\gamma/(1 - \exp(-2\gamma h))$ under the alternative hypothesis. This implies that u_{ht} and v_{ht} are N(0, 1) (in (2.3) and (2.4)) and y(0) is $N(0, (1 - \exp(-2\gamma h))^{-1})$ in (2.3). The experiment was performed for an alternative hypothesis specified as $\gamma = 0.2$ and the following grid of values for both the span S and the number of observations T: 8, 16, 32, 64, 128, 256, 512. The power against alternatives other than $\gamma = 0.2$ can be obtained from the tables. Since $\beta_h = \exp(-\gamma h)$, one can read the power of a test against the alternative that $\gamma = 0.2*2^j$ by reading j rows down.

The row "inf" gives the power of a test in the limiting case where the span tends to infinity. Since $\beta_h = \exp(-\gamma S/T)$, we have $\beta \to 0$ as $S \to \infty$ and the values reported in 'row inf' are simply the powers of a test that $\beta = 1$ against an alternative that $\beta = 0$. The row "C.V." lists the critical values used under the null hypothesis of a unit root. Note that under H_0, the distributions of the statistics are invariant with respect to h, the sampling interval; hence for a given value of T, the same critical values apply across values of S, the span. Finally, the row labelled 'size' gives the estimated exact size of the test obtained using again 2 000 replications by simulating the model under the null hypothesis of a random walk. The standard error of each entry is, in this case, 0.005.

5. ANALYSIS OF THE RESULTS

Table I presents the results concerning the normalized least-squares coefficient in a first-order autoregression using the level of the series $\{y_t\}$. The results are similar to those obtained in Shiller and Perron (1985) which considers the coefficient in a regression that includes a constant term. Here the power is slightly higher since some variability is excluded due to the fact that the mean of the series is not estimated. To summarize, with a small span the power is low and does not significantly increase as T increases. As the span increases, the power increases significantly. For any given span, the power appears to converge to a limit between 0 and 1, this limit increasing with the span. The results for the t-statistic (Table II) are similar except for the fact that the power is somewhat higher (this is contrary to the results obtained when a constant is included in the regression, see Shiller and Perron (1985)).

Table III presents the results concerning the locally best invariant test statistic proposed by King. As expected from a locally best invariant test the power for small spans (which correspond to higher frequencies for a given T and therefore alternatives closer to the null) is higher than for most other tests. This is verified for span of 8 and 16 units. Nevertheless the power remains rather low. For spans greater than 16 the power of King's test gets comparatively worse than the power of the previous two tests. As is the case for the statistics analysed previously, the power increases in both directions i.e. as T increases and as S increases. The power seems to level to some limit as T increases for a given span. This limit is lower than for the previous two tests for spans greater than 16.

Table IV presents the results concerning the uniformly most powerful invariant test proposed by Bhargava. The results are very similar to the ones with the OLS coefficient and its t-statistic with the power marginally greater for low spans and marginally smaller for large spans. Bhargava's test appears to provide an interesting alternative to the serial correlation coefficient.

The tests analysed so far are the only ones for which the power does not eventually decrease as T increases with a given fixed span. It is noteworthy that all the following tests interestingly share this feature.

Table V presents the results for the von Neumann test. Compared to the other tests the power is much lower, e.g., 0.11 for a span as large as 64 and any value of T. For a given span the power increases until T = 64 and then decreases significantly. Note that the test is significantly biased for low values of T and S. Although not shown here, it was found that the powers for a left-sided test (the explosive side) are significantly different from 0 for spans up to 64 indicating that for these values the test does not discriminate the direction of the alternative (see Perron (1986)). The results for the first-order correlation coefficient of the first-differences are presented in Table VI. The same comments apply as for the von Neumann statistic. The power here is maximized when T = 32 for all values of S except when S = 512 for which it is maximized at T = 64.

The turning point test (Table VII) has very low power, a maximum of 0.224 is obtained for S = 512 and T = 64 but does not appear to be biased for any pairs of values of T and S. It is clear in this case that the power tends to the size of the test as T increases for any given fixed span. Here the power is maximized at a value of T between 8 and 32 depending on the magnitude of the span. The Wald-Wolfowitz test statistic (Table VIII) and the rank correlation coefficient (Table IX) are appreciably more powerful and do not appear biased. Again the power eventually decreases as T increases for a given fixed span and is maximized at T = 32 or 64. The same comments apply to the rank version of the Von Neumann ratio proposed by Bartel (Table X).

The test statistics based on the ranks of the series $\{\Delta y_t \Delta y_{t-1}\}$ are presented in Table XI (sign test), Table XII (Wilcoxon signed-rank test) and Table XIII (Van der Waerden test statistic). These tests have very low power up to a span of 64 units for all values of T. The power is noticeably higher for larger spans. It is maximized at T = 32 for any given fixed span except for a span of 512 where it is maximized at T = 64. Again the power seems to converge towards the size of the test as T increases for any fixed value of S.

Several interesting features emerge from this simulation experiment. First, one can compare the various tests in terms of their power properties. Among the class of non-parametric tests analysed, the Wald-Wolfowitz statistic performs best, closely followed by the Wilcoxon signed-rank test and the Van der Waerden statistic. The Wald-Wolfowitz statistic seems almost as good as the parametric test based on the first-order correlation coefficient of the first-differences and better than the Von Neumann ratio. However, the best tests are those based on the original series $\{y_t\}$: the normalized least squares coefficient in a first-order autoregression and its t-statistic and Bhargava's test.

However, the most interesting feature emerging from the simulation experiment is the behavior of the power function of the tests as the sampling interval is varied. The common pattern is that all tests of the random walk based on a test of randomness in the first-differenced series have a power that eventually declines as the sampling interval decreases. Furthermore, the evidence for low spans suggests that the limit of the power as the sampling interval converges to zero is the size of the test. On the other hand, tests based on the original series $\{y_t\}$ have power that increases as more observations are added keeping a fixed span, though the power tends to level off as the sampling interval converges to zero.

This power property has interesting implications. First, when the power function converges to the size of the test as the sampling interval converges to zero, more observations are not necessarily better and too many may yield useless tests if they are concentrated in a data set with a small span. Indeed, it may be the case that higher power can be achieved by throwing away observations if the span is kept fixed, for example, by going from monthly to quarterly observations. Secondly, for test statistics having this property, there appears to be an optimal sampling interval (for any given fixed span) that maximizes the power of the test. This number of observations appears, from the simulation results, to be relatively small. This is not the case when considering tests based on the original series $\{y_t\}$. Here more observations are always better, though the marginal contribution of each addition is marginal if the span is fixed.

The feature that is common to both classes of tests is that the power is much more influenced by the span of the data than by the number of observations per se. While each increases in T, the sample size, has at most a marginally declining effect, the power appears to converge rapidly to one if both the span and the number of observations are increased.

In the notation of Perron (1987), the results suggested by the experiments are that tests based on the original series $\{y_t\}$ are consistent as long as the span increases with the sample size and this at any positive rate. Equivalently, they are consistent as long as the sampling interval does not converge to zero at a rate greater or equal to T (recall that $T = S/h$). On the other hand, tests of the random walk based on a test of randomness on the first-differences of the data are consistent only if the span increases with T at a higher rate. From the arguments presented in the next section, this rate is $T^{\frac{1}{2}}$, i.e. the span must increase at least at rate $T^{\frac{1}{2}}$ for the tests to be consistent. These empirical results are justified theoretically in the next section.

It must be stressed that the conclusions reached apply to the particular setting considered, i.e. testing the random walk hypothesis against an alternative that the process is stationary. The properties described above may not hold against other alternatives such as a random walk with correlated residuals. However, the results appear to be robust to other specifications concerning the distribution of the errors (see Perron (1986), chap. 1).

6. A THEORETICAL ANALYSIS OF THE CONSISTENCY PROPERTY

The results described in Perron (1987) can be used to analyse the consistency properties of the tests described here. The presentation is informal and more details on the method can be found in that paper. For ease of presentation, we analyse the special case where the initial observation is set at 0, i.e. $y(0) = 0$. The same conclusions follow letting $y(0)$ have an arbitrary pre-specified distribution including the one described in Section 2.

To analyse the limiting distribution of the statistics, we consider the process (2.3) embedded in a triangular array which allows the sampling interval h and the sample size T to be related as T increases. Each variable is then indexed by, say, n such that $T_n = S_n/h_n$. We require $T_n \to \infty$ and $h_n \to 0$ as $n \to \infty$. We can then define a triangular array of random variables $\{\{y_{nt}\}_{t=1}^{T_n}\}_{n=1}^{\infty}$. For a given n, the sequence $\{y_{nt}\}_{t=1}^{T_n}$ is generated by

$$y_{nt} = \exp(-\gamma h_n)y_{nt-1} + u_{nt} \qquad t = 1, \ldots, T_n \qquad (6.1)$$

where the innovation sequence $\{u_{nt}\}_{t=1}^{T_n}$ is i.i.d. normal with mean 0 and variance $\sigma^2(1 - \exp(-2\gamma h_n))/2\gamma$, and $y(0) = 0$.

The following lemma, proved in Perron (1987), describes the limiting distribution of the sample moment of $\{y_{nt}\}$ as $T_n \to \infty$ with the span S fixed for all n. The proof uses methods originally introduced in a series of papers by Phillips concerning continuous records asymptotic and near-integrated systems [see Phillips (1987a, 1987b, 1988)].

<u>Lemma 1</u>: If $\{\{y_{nt}\}_1^{T_n}\}_1^{\infty}$ is a triangular array of random variables defined by (6.1) and $y(0) = 0$, then as $n \to \infty$ and $h_n \to 0$ with $S_n = S$ for all n:

a) $\quad y_{T_n} \to S^{\frac{1}{2}}\sigma J_c(1)$

b) $\quad h_n\Sigma_1^{T_n} y_{nt} \to S^{3/2}\sigma\int_0^1 J_c(r)dr$

c) $h_n \Sigma_{t=1}^{T_n} y_{nt}^2 \rightarrow S^2\sigma^2 \int_0^1 J_c(r)^2 dr$

d) $\Sigma_{t=1}^{T_n} y_{nt-1}u_{nt} \rightarrow S\sigma^2 \int_0^1 J_c(r)dw(r)$

e) $\Sigma_{t=1}^{T_n} u_{nt}^2 \rightarrow S\sigma^2$

where $J_c(r) = \int_0^r e^{(r-s)c} dw(s)$ and $c = \gamma S$.

The statistics described in Section 3 can be written in terms of the triangular array of random variables defined by (6.1). These depend only on the sample moments defined in Lemma 1. The results in Lemma 1 can be used directly to derive the limiting distribution of these statistics both under the null ($\gamma = 0$) and alternative ($\gamma > 0$) hypotheses. The statistics are now indexed by the subscript n to emphasize that they are analysed under the triangular array of random variables defined by (6.1).

Theorem 1: If $\{\{y_{nt}\}_1^{T_n}\}_1^\infty$ is a triangular array of random variables defined by (6.1) and $y(0) = 0$, then as $n \rightarrow \infty$ and $h_n \rightarrow 0$ with S fixed:

a) $T_n(\hat{\beta}_n - 1) \rightarrow c + \{\int_0^1 J_c(r)^2 dr\}^{-1}\{\int_0^1 J_c(r)dw(r)\}$

b) $t_{\hat{\beta}_n} \rightarrow [c + \{\int_0^1 J_c(r)^2 dr\}^{-1}\{\int_0^1 J_c(r)dw(r)\}][\int_0^1 J_c(r)^2 dr]^{\frac{1}{2}}$

c) $d_n \rightarrow J_c(1)^2$

d) $T_n R_n \rightarrow [\int_0^1 J_c(r)^2 dr - (\int_0^1 J_c(r)dr)^2]^{-1}$

where $c = \gamma S$ and $J_c(r) = \int_0^r e^{(r-s)c} dw(s)$.

As a corrolary to Theorem 1, we can obtain the limiting distribution of the various statistics under the null hypothesis of a random walk. To do this, we simply let $\gamma = 0$ and hence $c = 0$ and $J_c(r) = w(r)$.

Corrolary 1: If $\{\{y_{nt}\}_1^{T_n}\}_1^\infty$ is a triangular array of random variables defined by (6.1) with $\gamma = 0$ and $y(0) = 0$, then as $n \to \infty$ and $h_n \to 0$ with S fixed:

a) $$T_n(\hat{\beta}_n - 1) \to [\int_0^1 w(r)^2 dr]^{-1} \int_0^1 w(r)dw(r)$$

$$= [\int_0^1 w(r)^2 dr]^{-1}(\tfrac{1}{2})(w(1)^2 - 1)$$

b) $$t_{\hat{\beta}_n} \to [\int_0^1 w(r)^2 dr]^{-\frac{1}{2}}(\tfrac{1}{2})(w(1)^2 - 1)$$

c) $$d_n \to w(1)^2$$

d) $$T_n R_n \to [\int_0^1 w(r)^2 dr - (\int_0^1 w(r)dr)^2]^{-1}.$$

It can be shown that the results of Corrolary 1 apply under any path for the sampling interval (see Perron (1987)). The same is not true for the limiting distribution under the alternative as given in Theorem 1. A few interesting results emerge from Theorem 1 and Corrolary 1. First, the limiting distributions of the statistics are non-degenerate and finite as $h \to 0$ keeping a fixed span. This proves the conjecture from the Monte Carlo experiment to the effect that the power function of tests based on the original series $\{y_t\}$ have a non-degenerate limiting power function as $h \to 0$. That is, the power does not converge to 1 but to some value which depends on the span. Some exact values were obtained for the statistic $T(\hat{\beta} - 1)$ in Perron (1987).

One can also easily obtain the limiting power function as both the sampling interval and the span converge to zero. Since $c = \gamma S$, the limiting distribution of the statistics as $S \to 0$ are obtained by taking the limit as $c \to 0$. This implies that the limiting distributions under the alternative are the same as the limiting distributions under the null (see Corrolary 1); hence, the power function converges to the size of the tests as $S \to 0$ with $T \to \infty$, i.e. when the sampling interval converges to zero at a rate faster than T.

Arguments similar to those given in Perron (1987) can be used to show that the statistics considered in Theorem 1 all converge to $-\infty$, under the alternative, if $S \to \infty$ as $T \to \infty$. Hence, if the span is increasing (at whatever rate) with the sample size, the test statistics are consistent. This confirms the conjecture from the Monte Carlo study.

When considering tests of the random walk hypothesis based on a test of randomness on the first-differenced data, a different approach must be taken. We shall, however, not be as precise as above. Our argument is based on the results of Theorem 2 below which shows that for a well-defined class of test statistics, we obtain consistency if and only if the span is increasing at a rate greater than $T^{\frac{1}{2}}$. This class of statistics has a root T convergence rate and satisfies the so-called Pitman's conditions for a non-degenerate local asymptotic power function with contiguous alternatives that approach the null at a rate $T^{\frac{1}{2}}$. Our claim is that the tests presented here that use the first-differenced data satisfy those conditions. To be more precise, we would have to verify on a case-by-case basis whether or not they apply. For the moment, this is a conjecture except for the statistics based on the first-order correlation coefficient where it was formally shown in Perron (1987) that the conditions were satisfied.

<u>Theorem 2</u>: Let J_n be a test statistic based on the first T_n observations sampled at intervals of length h_n, and let the critical region be $J_n \geq \lambda_n$. Suppose that:

a) $\lim_{n\to\infty} P(J_n \geq \lambda_n) = \alpha > 0.$

b) there exist functions $\mu(\theta)$ and $\sigma(\theta)$ such that

$$\lim_{n\to\infty} P\left\{ T_n^{\frac{1}{2}} \frac{J_n - \mu(\theta_0 + \delta T_n^{-\frac{1}{2}})}{\sigma(\theta_0 + \delta T_n^{-\frac{1}{2}})} < y \,\Big|\, \theta_0 + \delta T_n^{-\frac{1}{2}} \right\} = \Phi(y)$$

for every real y, where $\Phi(y)$ is the distribution function of $N(0,1)$.

c) $\mu(\theta)$ has a derivative $\mu'(\theta_0)$ at $\theta = \theta_0$, which is positive and $\sigma(\theta)$ is continuous at θ_0.

Then J_n is a consistent test of the random walk hypothesis against stationary alternatives if and only if $h_n = 0(T_n^a)$ for any $a < \frac{1}{2}$.

The conditions of Theorem 2 are usually referred to as Pitman's conditions. They are related to the concept of the Pitman efficiency which considers the local asymptotic power of test statistics under contiguous alternatives of the form $\theta_0 + \delta T_n^{-\frac{1}{2}}$ (for details see, e.g., Rao (1973), section 7.a.7).

The proof of the theorem is presented in the appendix. To get some intuition about the result, note first that we can write $\theta_n = \beta_n = \exp(-\gamma h_n)$ and therefore $\theta_n = \theta_0 + T_n^{-\frac{1}{2}}\delta_n$ where $\delta_n = T_n^{\frac{1}{2}}(\exp(-\gamma h_n) - 1)$ and $\theta_0 = 1$. Given condition (b) of the theorem, the statistic J_n has a non-degenerate limiting distribution against local alternatives that approach the null value at a rate of $T_n^{\frac{1}{2}}$. If the alternative value approaches the null at a rate lower than $T_n^{\frac{1}{2}}$, the power function converges to 1. However, if it approaches the null at a rate faster than $T_n^{\frac{1}{2}}$, the power converges to the size of the test. When the sampling interval is varied, the value of the autoregressive parameter changes. When the sampling interval decreases, β_n converges to 1 at a rate which depends on the rate at which h the sampling interval converges to 0 relative to the sample size. For instance, if the span of the data is fixed, the sampling interval converges to 0 at the same rate as the sample size converges to infinity. Hence, in this case we expect the test to be inconsistent. This is proved formaly in Theorem 2. In general, test statistics which satisfy the conditions of Theorem 2 will be consistent tests of the random walk hypothesis against stationary alternatives only if the sampling interval does not decrease at a rate greater or equal to $T^{\frac{1}{2}}$.

Of course, in order for this theorem to explain the empirical findings suggested by the Monte Carlo experiments, it must be the case that the statistics analysed satisfy the conditions of Theorem 2. This was shown to be the case for the first-order autocorrelation coefficient applied to the first-differenced series in Perron (1987). For the other statistics, it remains a plausible conjecture which should be verified on a case-by-case basis. However, it must be understood that a great many statistics satisfy those conditions. Most test statistics are usually asymptotically normal with a root T convergence rate. If this holds uniformly in a neighborhood of the null hypothesis, then condition (b) will be satisfied. Indeed, see Rao (1973, p. 468), a sufficient condition for (b) to hold is that

$$\lim_{T\to\infty} P\left\{ T^{\frac{1}{2}} \frac{J_n - \mu(\theta)}{\sigma(\theta)} < y \,\middle|\, \theta \right\} = \Phi(y) \tag{6.2}$$

uniformly in θ for $\theta_0 \leqslant \theta \leqslant \theta_0 + \eta$ where η is any positive number. Any test that satisfies (6.2) and conditions (a) and (c) will behave according to the results in Theorem 2. The conjecture, not verified explicitly here, is that the statistics based on the first-differences of the data considered in this study satisfy these conditions.

7. CONCLUSIONS

This paper has presented an extended analysis of the behavior of the power function for a wide class of tests of the random walk hypothesis against stationary first-order autoregressive alternatives.

Several features that emerged are worth mentioning. First, the power depends more importantly on the span of the data rather than the number of observations per se for all statistics considered. It is preferable to have a large span of data even, in most cases, if this entails a smaller number of observations available. Second, there is a notable difference between tests using the original level of the series and tests based on testing for randomness in first-differenced data. In the latter case too many observations, for a given fixed span, may destroy the power. It was shown, in particular, that as the number of observations increases keeping a fixed span, the power converges to the size of the test. It may be the case that higher power can be obtained by deleting observations while keeping the span fixed. This feature is not present for the class of tests based on the original undifferenced series. In that case, more observations always lead to higher power though the marginal contribution of each additional observation is quickly declining. Finally, when comparing the power of the various tests, our results suggest that the Dickey-Fuller type procedure and Barghava's test stand out as the preferred tests.

These results relate interestingly to the concept of consistency of a testing procedure. The usual consistency criterion states that a test is consistent if the power function converges to one as the number of observations increases to infinity for any given fixed alternative. The requirement of a fixed alternative has led, in a time series context, to analyse the asymptotic behavior with a fixed sampling interval as T increases. However, one can view the fixed alternative in terms of the parameters of the continuous time model and, in this case, there is no need to consider the sampling interval fixed as the sample size increases. As shown here (and detailed in Perron (1987)), a richer set of properties concerning the behavior of the power function can be obtained by considering a continuum of consistency criteria indexed by the possible paths of the sampling interval as T increases. What has been shown here is that tests of the random walk hypothesis ($\gamma = 0$) based on the original series are consistent against a fixed mean-reverting alternative ($\gamma < 0$) if and only if the sampling interval does not decrease at a rate faster or equal to T; or, equivalently, if and only if the span is increasing as the sample size increases. On the other hand, tests based on the differenced series are consistent if and only if the span increases at least at rate $T^{\frac{1}{2}}$. Therefore in a well-defined sense, the former class of tests can be said to dominate the latter since they are consistent over a wider range of possible paths for the sampling interval.

The Monte Carlo experiment showed how these consistency properties are valuable in providing information on the behavior of the power of the tests in finite samples. Of course, the results presented here need not carry over to different models. Nevertheless they show that in a time series context the notion that more observations is desirable clearly depends not only on the statistics and the null and alternative hypotheses considered but also on the time interval between each observation.

APPENDIX

Proof of Theorem 1

Part (a) is proved in Perron (1987). To prove part (b), the t-statistic should be written as:

$$t_{\hat{\beta}_n} = T_n(\hat{\beta}_n - 1)(h_n S^{-1} \Sigma_1^{T_n} y_{nt-1}^2)^{\frac{1}{2}} / T_n^{\frac{1}{2}} \hat{\sigma}_n.$$

Now

$$T_n \hat{\sigma}_n^2 = \Sigma_1^{T_n}(y_{nt} - \hat{\beta}_n y_{nt-1})^2$$

$$= \Sigma_1^{T_n} u_{nt}^2 - 2(\hat{\beta}_n - \beta_n)\Sigma_1^{T_n} u_{nt} y_{nt-1} + (\hat{\beta}_n - \beta_n)^2 \ \Sigma_1^{T_n} y_{nt-1}^2.$$

Noting that $h_n^{-1}(\hat{\beta}_n - \beta_n) = 0(1)$ (see Perron (1987)) and using Lemma 1, we have $T_n \hat{\sigma}_n^2 \to S\sigma^2$ as $n \to \infty$ and $h_n \to 0$ with S fixed. Hence

$$t_{\hat{\beta}_n} \to \lfloor c + \int_0^1 J_c(r)dw(r)\{\int_0^1 J_c(r)^2 dr\}^{-1}\rfloor. \quad \lfloor S\sigma^2 \int_0^1 J_c(r)^2 dr\rfloor^{\frac{1}{2}} (S\sigma^2)^{\frac{1}{2}}$$

which proves part (b) upon simplification.

To prove part (c), we first note that

$$d_n = y_{T_n}^2 / \Sigma_1^{T_n}(y_{nt} - y_{nt-1})^2.$$

The denominator can be written as:

$$\Sigma_1^{T_n}(y_{nt} - y_{nt-1})^2 = (\beta_n - 1)^2 \Sigma_1^{T_n} y_{nt-1}^2 + 2(\beta_n - 1)\Sigma_1^{T_n} y_{nt-1} u_{nt} + \Sigma_1^{T_n} u_{nt}.$$

Since $h_n^{-1}(\beta_n - 1) = h_n^{-1}(\exp(-\gamma h_n) - 1) \to -\gamma$ as $h_n \to 0$, the first two terms vanish asymptotically and

$$\Sigma_1^{T_n}(y_{nt} - y_{nt-1})^2 \to S\sigma^2. \qquad \text{(A.1)}$$

Finally

$$d_n \to S\sigma^2 J_c(1)^2 / S\sigma^2 = J_c(1)^2$$

using Lemma 1. To prove part (d), $T_n R_n$ should be written as:

$$T_nR_n = \Sigma_{t=1}^{T_n}(y_{nt} - y_{nt-1})^2/T_n^{-1}\Sigma_{t=0}^{T_n}(y_{nt} - \overline{Y}_n)^2$$
$$= S\Sigma_1^{T_n}(y_{nt} - y_{nt-1})^2/[h_n\Sigma_1^{T_n}y_{nt}^2 - h_n^2S^{-1}(\Sigma_1^{T_n}y_{nt})^2] + o(1)$$
$$\to S^2\sigma^2/[S^2\sigma^2\int_0^1 J_c(r)^2dr - S^2\sigma^2(\int_0^1 J_c(r)dr)^2]$$
$$= [\int_0^1 J_c(r)^2dr - (\int_0^1 J_c(r)dr)^2]^{-1} \text{ using (A.1) and Lemma 1.}$$

Proof of Theorem 2

The proof is closely related to the development in Rao (1973), section 7.a.7. Here $\theta = \beta_n = \exp(-\gamma h_n)$ and $\theta_0 = 1$ since under the null hypothesis $\gamma = 0$. Hence, under the alternative hypothesis we can write $\theta = \theta_0 + (\exp(-\gamma h_n) - 1) = \theta_0 + \delta_n T_n^{-\frac{1}{2}}$ where $\delta_n = T_n^{\frac{1}{2}}(\exp(-\gamma h_n) - 1)$. Now denote the power function of J_n by $P_n(\theta_0 + \delta_n T_n^{-\frac{1}{2}})$ under the alternative hypothesis that $\theta = \theta_0 + \delta_n T_n^{-\frac{1}{2}}$. By definition, we have:

$$P_n(\theta_0 + \delta_n T_n^{-\frac{1}{2}}) = P[J_n \geq \lambda_n \mid \theta_0 + \delta_n T_n^{-\frac{1}{2}}]$$
$$= P\left\{T_n^{\frac{1}{2}} \frac{J_n - \mu(\theta_0 + \delta_n T_n^{-\frac{1}{2}})}{\sigma(\theta_0 + \delta_n T_n^{-\frac{1}{2}})} \geq T_n^{\frac{1}{2}} \frac{\lambda_n - \mu(\theta_0 + \delta_n T_n^{-\frac{1}{2}})}{\sigma(\theta_0 + \delta_n T_n^{-\frac{1}{2}})} \;\middle|\; \theta_0 + \delta_n T_n^{-\frac{1}{2}}\right\}$$

$$= \Phi\left(-T_n^{\frac{1}{2}} \frac{\lambda_n - \mu(\theta_0 + \delta_n T_n^{-\frac{1}{2}})}{\sigma(\theta_0 + \delta_n T_n^{-\frac{1}{2}})}\right) + \varepsilon_n(\delta_n) \tag{A.2}$$

using condition (b), where the convergence is uniform in y, although not in δ_n. Here $\varepsilon_n(\delta_n) \to 0$ as $n \to \infty$. Substituting $\delta_n = 0$ we have

$$P_n(\theta_0) = \Phi\left(-T_n^{\frac{1}{2}} \frac{\lambda_n - \mu(\theta_0)}{\sigma(\theta_0)}\right) + \varepsilon_n(0). \tag{A.3}$$

Taking the limit of both sides of (A.3) and noting that $\lim_{n\to\infty} P_n(\theta_0) = \alpha$ using condition (a), we have $\alpha = \lim_{n\to\infty} \Phi(-T_n^{\frac{1}{2}}(\lambda_n - \mu(\theta_0))/\sigma(\theta_0))$ which shows that $-T_n^{\frac{1}{2}}(\lambda_n - \mu(\theta_0))/\sigma(\theta_0) = a + \eta_n$ with $\eta_n \to 0$ as $n \to \infty$ and $\alpha = \Phi(a)$. Therefore $\lambda_n = -T_n^{-\frac{1}{2}}(a + \eta_n)\sigma(\theta_0) + \mu(\theta_0)$. Hence, the argument of Φ in (A.2) has the following limit:

$$-T_n^{\frac{1}{2}}(\lambda_n - \mu(\theta_0 + \delta_n T_n^{-\frac{1}{2}}))/\sigma(\theta_0 + \delta_n T_n^{-\frac{1}{2}})$$
$$= [-\delta_n[\mu(\theta_0 + \delta_n T_n^{-\frac{1}{2}}) - \mu(\theta_0)]/\delta_n T_n^{-\frac{1}{2}} + (a + \eta_n)\sigma(\theta_0)]/\sigma(\theta_0 + \delta_n T_n^{-\frac{1}{2}})$$

$$= [-\delta_n \mu'(\theta_0) + a\ \sigma(\theta_0)]/\sigma(\theta_0) + \varepsilon'_n$$

$$= a - \delta_n \mu'(\theta_0)/\sigma(\theta_0) + \varepsilon'_n$$

where $\varepsilon'_n \to 0$ as $n \to \infty$. Hence

$$\lim_{n\to\infty} P_n(\theta_0 + \delta_n T_n^{-\frac{1}{2}}) = \Phi[\lim_{n\to\infty}(a - \delta_n \mu'(\theta_0)/\sigma(\theta_0))].$$

Now, $\delta_n = T_n^{\frac{1}{2}}(\exp(-\gamma h_n) - 1)$ and we have: i) $\delta_n \to 0$ if $h_n = 0(T_n^a)$ for any $a > \frac{1}{2}$; ii) $\delta_n \to -\infty$ if $h_n = 0(T_n^a)$ for any $a < \frac{1}{2}$ and iii) $\delta_n \to -\gamma$ if $h_n = 0(T_n^{\frac{1}{2}})$.

Finally, this shows that

$$\lim_{n\to\infty} P_n(\theta_0 + \delta_n T_n^{-\frac{1}{2}}) = \begin{cases} \Phi(a) = \alpha & \text{if} \quad h_n = 0(T_n^a) \text{ for any } a > \frac{1}{2} \\ \Phi(\infty) = 1 & \text{if} \quad h_n = 0(T_n^a) \text{ for any } a < \frac{1}{2} \\ \Phi[a + \gamma\mu'(\theta_0)/\sigma(\theta_0)] & \text{if} \quad h_n = 0(T_n^{\frac{1}{2}}). \end{cases}$$

Table I
Normalized OLS Coefficient

S \ T	8	16	32	64	128	256	512
8	0.078	0.077	0.080	0.069	0.069	0.071	0.075
16	0.148	0.158	0.166	0.155	0.166	0.155	0.155
32	0.303	0.359	0.407	0.413	0.430	0.431	0.421
64	0.519	0.740	0.839	0.868	0.903	0.903	0.908
128	0.670	0.947	0.996	1.000	1.000	1.000	1.000
256	0.704	0.988	1.000	1.000	1.000	1.000	1.000
512	0.705	0.992	1.000	1.000	1.000	1.000	1.000
INF	0.705	0.992	1.000	1.000	1.000	1.000	1.000
C.V.	-6.275	-7.068	-7.497	-7.795	-7.760	-7.943	-8.079
Size	0.057	0.053	0.050	0.051	0.057	0.053	0.053

Table II
t-Statistic on $\hat{\beta}$

S \ T	8	16	32	64	128	256	512
8	0.150	0.142	0.141	0.132	0.116	0.122	0.132
16	0.232	0.225	0.229	0.224	0.216	0.212	0.222
32	0.396	0.436	0.488	0.480	0.463	0.473	0.477
64	0.609	0.778	0.863	0.892	0.903	0.912	0.913
128	0.737	0.948	0.995	0.999	0.999	1.000	1.000
256	0.769	0.987	1.000	1.000	1.000	1.000	1.000
512	0.770	0.990	1.000	1.000	1.000	1.000	1.000
INF	0.770	0.990	1.000	1.000	1.000	1.000	1.000
C.V.	-1.972	-1.974	-1.946	-1.934	-1.959	-1.957	-1.946
Size	0.067	0.060	0.054	0.050	0.056	0.047	0.056

Table III
Locally Best Invariant

S \ T	8	16	32	64	128	256	512
8	0.085	0.095	0.085	0.098	0.093	0.099	0.091
16	0.120	0.131	0.127	0.120	0.117	0.124	0.138
32	0.158	0.168	0.166	0.168	0.161	0.179	0.174
64	0.186	0.207	0.211	0.234	0.236	0.248	0.244
128	0.199	0.268	0.280	0.325	0.345	0.335	0.330
256	0.207	0.263	0.350	0.410	0.430	0.473	0.460
512	0.214	0.286	0.382	0.469	0.537	0.593	0.605
INF	0.214	0.286	0.385	0.515	0.681	0.857	0.948
C.V. $\times 10^5$	60.300	27.098	12.890	6.291	3.108	1.545	0.770
Size	0.047	0.046	0.045	0.050	0.053	0.049	0.050

Table IV
Uniformly Most Powerful Invariant

S \ T	8	16	32	64	128	256	512
8	0.059	0.065	0.067	0.075	0.074	0.074	0.053
16	0.097	0.094	0.106	0.099	0.088	0.110	0.101
32	0.163	0.183	0.215	0.234	0.234	0.238	0.225
64	0.290	0.455	0.547	0.599	0.621	0.641	0.626
128	0.400	0.770	0.932	0.987	0.991	0.991	0.992
256	0.444	0.901	1.000	1.000	1.000	1.000	1.000
512	0.427	0.921	1.000	1.000	1.000	1.000	1.000
INF	0.427	0.923	1.000	1.000	1.000	1.000	1.000
C.V.	2.115	1.324	0.749	0.400	0.207	0.105	0.055
Size	0.058	0.057	0.062	0.054	0.049	0.055	0.051

Table V
The von Neumann Ratio

S \ T	8	16	32	64	128	256	512
8	0.053	0.050	0.054	0.051	0.051	0.053	0.048
16	0.054	0.075	0.064	0.073	0.056	0.064	0.059
32	0.096	0.090	0.086	0.084	0.078	0.062	0.063
64	0.149	0.175	0.160	0.150	0.110	0.090	0.077
128	0.194	0.327	0.360	0.306	0.233	0.157	0.127
256	0.220	0.498	0.682	0.691	0.555	0.396	0.255
512	0.211	0.513	0.854	0.950	0.949	0.851	0.667
INF	0.218	0.497	0.888	0.997	1.000	1.000	1.000
C.V.	3.450	2.958	2.647	2.433	2.307	2.214	2.149
Size	0.047	0.059	0.051	0.058	0.054	0.058	0.058

Table VI
First-order Correlation Coefficient

S \ T	8	16	32	64	128	256	512
8	0.039	0.046	0.067	0.055	0.058	0.064	0.055
16	0.050	0.060	0.086	0.064	0.066	0.068	0.058
32	0.082	0.111	0.116	0.089	0.081	0.074	0.067
64	0.158	0.224	0.200	0.161	0.132	0.107	0.089
128	0.225	0.416	0.430	0.343	0.257	0.199	0.139
256	0.250	0.582	0.766	0.718	0.581	0.412	0.280
512	0.251	0.620	0.900	0.965	0.953	0.854	0.650
INF	0.251	0.622	0.928	0.997	1.000	1.000	1.000
C.V.	-1.645	-1.645	-1.645	-1.645	-1.645	-1.645	-1.645
Size	0.027	0.039	0.056	0.047	0.051	0.061	0.053

Table VII
Turning Point Test

S \ T	8	16	32	64	128	256	512
8	0.073	0.073	0.061	0.050	0.048	0.048	0.045
16	0.077	0.078	0.061	0.050	0.047	0.048	0.044
32	0.090	0.085	0.071	0.053	0.048	0.045	0.044
64	0.112	0.102	0.080	0.054	0.049	0.043	0.045
128	0.134	0.139	0.098	0.065	0.058	0.049	0.046
256	0.141	0.184	0.160	0.110	0.083	0.058	0.046
512	0.141	0.194	0.224	0.216	0.151	0.088	0.062
INF	0.141	0.194	0.240	0.363	0.568	0.812	0.978
C.V.	1.645	1.645	1.645	1.645	1.645	1.645	1.645
Size	0.069	0.069	0.057	0.048	0.047	0.046	0.045

Table VIII
Wald-Wolfowitz Statistic

S \ T	8	16	32	64	128	256	512
8	0.047	0.051	0.063	0.054	0.052	0.061	0.055
16	0.054	0.054	0.070	0.060	0.058	0.064	0.056
32	0.073	0.089	0.095	0.076	0.074	0.072	0.065
64	0.119	0.166	0.173	0.142	0.111	0.099	0.085
128	0.167	0.326	0.363	0.308	0.231	0.183	0.132
256	0.184	0.482	0.691	0.672	0.541	0.385	0.270
512	0.184	0.524	0.861	0.947	0.945	0.843	0.633
INF	0.184	0.526	0.889	0.996	1.000	1.000	1.000
C.V.	-1.645	-1.645	-1.645	-1.645	-1.645	-1.645	-1.645
Size	0.045	0.048	0.060	0.050	0.050	0.060	0.054

Table IX
Rank Correlation Coefficient

S \ T	8	16	32	64	128	256	512
8	0.029	0.037	0.048	0.047	0.056	0.062	0.051
16	0.031	0.044	0.053	0.053	0.061	0.067	0.053
32	0.044	0.061	0.082	0.076	0.071	0.074	0.061
64	0.067	0.132	0.146	0.128	0.108	0.101	0.077
128	0.093	0.253	0.321	0.265	0.212	0.164	0.119
256	0.104	0.384	0.619	0.616	0.491	0.354	0.249
512	0.104	0.429	0.809	0.928	0.923	0.811	0.611
INF	0.104	0.430	0.849	0.992	1.000	1.000	1.000
C.V.	-1.645	-1.645	-1.645	-1.645	-1.645	-1.645	-1.645
Size	0.024	0.034	0.048	0.046	0.056	0.062	0.051

Table X
Rank Version of von Neumann Ratio

S \ T	8	16	32	64	128	256	512
8	0.054	0.045	0.055	0.049	0.059	0.062	0.052
16	0.069	0.054	0.064	0.054	0.062	0.067	0.053
32	0.090	0.080	0.092	0.079	0.074	0.074	0.061
64	0.141	0.161	0.158	0.132	0.118	0.101	0.076
128	0.188	0.296	0.335	0.274	0.213	0.165	0.120
256	0.199	0.429	0.644	0.620	0.496	0.356	0.249
512	0.201	0.472	0.811	0.928	0.922	0.808	0.610
INF	0.201	0.473	0.853	0.993	1.000	1.000	1.000
C.V.	1.645	1.645	1.645	1.645	1.645	1.645	1.645
Size	0.049	0.041	0.053	0.046	0.059	0.062	0.050

Table XI
Sign Test

S \ T	8	16	32	64	128	256	512
8	0.069	0.060	0.038	0.043	0.059	0.060	0.046
16	0.086	0.075	0.044	0.048	0.062	0.062	0.043
32	0.118	0.101	0.063	0.054	0.071	0.070	0.053
64	0.169	0.157	0.092	0.080	0.094	0.089	0.062
128	0.208	0.247	0.175	0.151	0.152	0.120	0.090
256	0.219	0.336	0.334	0.340	0.309	0.224	0.154
512	0.219	0.359	0.476	0.621	0.669	0.515	0.345
INF	0.219	0.360	0.517	0.844	0.995	1.000	1.000
C.V.	-1.645	-1.645	-1.645	-1.645	-1.645	-1.645	-1.645
Size	0.056	0.054	0.038	0.040	0.059	0.053	0.040

Table XII
Wilcoxon Signed-rank Test

S \ T	8	16	32	64	128	256	512
8	0.070	0.060	0.057	0.055	0.064	0.062	0.054
16	0.083	0.078	0.069	0.064	0.068	0.066	0.057
32	0.123	0.109	0.099	0.085	0.081	0.074	0.062
64	0.191	0.208	0.171	0.134	0.117	0.100	0.076
128	0.266	0.355	0.344	0.276	0.209	0.160	0.120
256	0.282	0.506	0.634	0.595	0.461	0.327	0.237
512	0.282	0.538	0.806	0.892	0.883	0.751	0.540
INF	0.282	0.539	0.842	0.987	1.000	1.000	1.000
C.V.	-1.645	-1.645	-1.645	-1.645	-1.645	-1.645	-1.645
Size	0.050	0.048	0.058	0.045	0.057	0.055	0.053

Table XIII
Van der Waerden Test Statistic

S \ T	8	16	32	64	128	256	512
8	0.070	0.054	0.058	0.058	0.065	0.064	0.055
16	0.083	0.073	0.071	0.068	0.072	0.068	0.057
32	0.123	0.107	0.110	0.089	0.085	0.080	0.063
64	0.191	0.212	0.183	0.147	0.124	0.106	0.080
128	0.266	0.366	0.370	0.304	0.225	0.183	0.127
256	0.282	0.521	0.672	0.651	0.512	0.372	0.265
512	0.282	0.557	0.841	0.924	0.927	0.817	0.615
INF	0.282	0.562	0.871	0.995	1.000	1.000	1.000
C.V.	-1.645	-1.645	-1.645	-1.645	-1.645	-1.645	-1.645
Size	0.050	0.045	0.055	0.048	0.056	0.059	0.052

REFERENCES

Anderson, R.L.: 1942, 'Distribution of the Serial Correlation Coefficient', Annals of Mathematical Statistics, 13, 1-13.

Arnold, L.: 1974, Stochastic Differential Equations: Theory and Applications, John Wiley and Sons, New York.

Bartels, R.: 1982, 'The Rank Version of Neumann's Ratio Test for Randomness', Journal of the American Statistical Association, 77, 40-46.

Bergström, A.R.: 1984, 'Continuous Time Stochastic Models and Issues of Aggregation Over Time', Ch. 20, pp. 1145-1212 in M.D. Intriligator and Z. Griliches (eds.), Handbook of Econometrics, 2, North-Holland, Amsterdam.

Bhargava, Alok: 1986, 'On the Theory of Testing for Unit Roots in Observed Time Series', Review of Economic Studies, 53, 369-384.

Dickey, D.A.: 1976, 'Estimation and Hypothesis Testing for Nonstationary Time Series', Ph.D. Thesis, Iowa State University, Ames.

Dufour, J.-M.: 1981, 'Rank Tests for Serial Dependence', Journal of Time Series Analysis, 2, 117-126.

Dufour, J.M. and M.L. King: 1986, 'Optimal Invariant Test for the Autocorrelation Coefficient in a Linear Regression Model', mimeo, Université de Montréal.

Dufour, J.-M. and P. Perron: 1984, 'Extended Tables for the von Neumann Test of Independence', mimeo, Université de Montréal.

Fuller, Wayne A.: 1976, Introduction to Statistical Time Series, John Wiley and Sons, New York.

Girard, Michel: 1983, 'Comportement experimental de divers tests d'indépendance pour des contre-hypothèses issues des modèles de séries chronologiques', unpublished M.Sc. Thesis, Département de mathématiques et de statistique, Université de Montréal.

Kendall, M.G. and A. Stuart: 1976, The Advanced Theory of Statistics, Haffner Press, London, 3rd edition.

King, M.L.: 1981, 'Notes on Testing for Unit Roots', mimeo.

Knoke, J.: 1977, 'Testing for Randomness Against Autocorrelation: Alternative Tests', Biometrika, 64, 523-529.

Koerts, J. and A.P.J. Abrahamse: 1969, On the Theory and Applications of the General Linear Model, Rotterdam, Rotterdam University Press.

Lepage, Y. and H. Zeidan: 1981, 'L'indépendance d'une série d'observations', Annales des sciences mathématiques du Québec, 5, 169-184.

Mood, A.M.: 1940, 'The Distribution Theory of Runs', Annals of Mathematical Statistics, 11, 367-392.

Moran, P.A.P.: 1948, 'Some Theorems on Time Series II. The Significance of the Serial Correlation Coefficient', Biometrika, 35, 255-260.

Neumann, J. von: 1942, 'Distribution of the Ratio of the Mean Square Successive Difference to the Variance', Annals of Mathematical Statistics, 13, 367-395.

Perron, Pierre: 1986, 'Hypothesis Testing in Time Series Regressions with a Unit Root', Ph.D. Dissertation, Yale University.

Perron, Pierre: 1987, 'Test Consistency with Varying Sampling Interval', Cahier de recherche no 4187, C.R.D.E., Université de Montréal.

Phillips, P.C.B.: 1987a, 'Time Series Regression with Unit Roots', Econometrica, 55, 277-302.

Phillips, P.C.B.: 1987b, 'Towards a Unified Asymptotic Theory for Autoregression', Biometrika, 74, 535-47.

Phillips, P.C.B.: 1988, 'Regression Theory for Near-Integrated Time Series', forthcoming in Econometrica.

Rao, C.R.: 1973, Linear Statistical Inference and Its Applications, 2nd edition, John Wiley and Sons, New York.

Rao, M.M.: 1978, 'Asymptotic Distribution of an Estimator of the Boundary Parameter of an Unstable Process', Annals of Statistics, 6, 185-190.

Sargan, J.D. and A. Bhargava: 1983, 'Testing Residuals from Least Squares Regression for Being Generated by a Gaussian Random Walk', Econometrica, 51, 153-174.

Shiller, R.J. and P. Perron: 1985, 'Testing the Random Walk Hypothesis: Power Versus Frequency of Observation', Economics Letters, 18, 381-386.

Waerden, B.W. van der: 1952, 'Order Tests for the Two-sample Problem and Their Power', Indagationes Mathematicae, series A, 14, 453-458.

Wald, A. and J. Wolfowitz: 1943, 'An Exact Test for Randomness in the Non-parametric Case Based on Serial Correlation', Annals of Mathematical Statistics, 14, 378-388.