Testing for Drift in a Time Series

Fabio Busetti and Andrew Harvey

DAE Working Paper No. 0237

Testing for Drift in a Time Series

Fabio Busetti and Andrew Harvey Bank of Italy
and Faculty of Economics, Cambridge University

December 2, 2002

Abstract

The paper presents various tests for assessing whether a time series is subject to drift. We first consider departures from the null hypothesis of no drift against the alternative of a deterministic and/or a non-stationary stochastic drift with initial value zero. We show that the standard t-test on the mean of first differences achieves high power in both directions of the alternative hypothesis and it seems preferable to locally best invariant tests specifically designed to test against a non-stationary drift. The test may be modified, parametrically or nonparametrically to deal with serial correlation. Tests for the null hypothesis of a non-stationary drift are then examined. The simple t-statistic, now standardized by the square root of the sample size, is again a viable alternative, but this time there is no need to correct for serial correlation. We present the asymptotic distribution of the test, provide critical values and compare its performance with that of the standard augmented Dickey-Fuller test procedures. We show that the t-test does not suffer from the large size distortions of the augmented Dickey-Fuller test for cases in which the variance of the nonstationary drift, the signal, is small compared to that of the stationary part of the model. The use of the tests is illustrated with data on global warming and electricity consumption.

KEYWORDS: Cramér-von Mises distribution, locally best invariant test, stochastic trend, unit root, unobserved components.

JEL classification: C22, C52.

1. Introduction

The question of whether a time series exhibits drift is an important one. In other words, does the series show a steady upward or downward movement over time that can be extrapolated into the future? If no drift is present, first differences of the series have zero mean and the eventual forecast function is constant.

We begin by considering how to test the null hypothesis of no drift. If the drift is taken to be fixed, it may be estimated and a simple $t-$statistic formed. The more general hypothesis of a stochastic drift may be tested by a variant of the test proposed by Nyblom and Mäkeläinen (1983) in which it is assumed that there is no drift under the null hypothesis. The asymptotic distribution of the test statistic under the null hypothesis belongs to the Cramér-von Mises family and a rejection may be interpreted as indicating the presence of a drift that is either deterministic or stochastic. Section 2 of the paper analyses the local power of this test and compares it with the local power of the $t-$test. Both tests are consistent against the hypothesis of deterministic or/and non-stationary stochastic drift with zero starting value. When serial correlation is present, the tests can be implemented nonparametrically or by fitting a time series model.

Section 3 reverses the roles of the null and alternative hypotheses. A test can again be based on the $t-$statistic but this time the null is that a drift is present. The test is nonparametric since when the $t-$statistic is divided by the square root of the sample size it has an asymptotic distribution that does not depend on any nuisance parameters when the drift is nonstationary. In fact because of this asymptotic result, the test statistic is formulated more naturally as the inverse of the coefficient of variation for first differences. The performance of this test is evaluated under a number of different scenarios with various unit root tests used as a benchmark for comparison. Extensions of the tests to deal with seasonally unadjusted data are given in section 4, while section 5 provides empirical examples. The conclusions are presented in section 6.

2. Testing against Stochastic and Deterministic Drift

Consider the data generating process

$$\Delta y_t = \beta_t + \eta_t, \quad t = 1, ..., T \tag{2.1}$$

where η_t is a serially uncorrelated disturbance term with mean zero and variance σ_η^2, and β_t is a drift, or slope, component that follows a random walk

$$\beta_t = \beta_{t-1} + \zeta_t, \qquad t = 1, ..., T, \tag{2.2}$$

with initial value, β_0, and ζ_t is a serially uncorrelated disturbance term with mean zero and variance σ_ζ^2. The forecast function is a linear trend with slope b_T, where b_T is the estimator of β_T.

If the drift is assumed to be deterministic, that is $\sigma_\zeta^2 = 0$, and η_t is normally distributed and serially independent, that is $\eta_t \sim NID(0, \sigma_\eta^2)$, the $t-$statistic for testing the null hypothesis that $\beta = 0$ is

$$t_\beta = T^{1/2}\widehat{\beta}/\widehat{\sigma} = T^{-1/2}(y_T - y_0)/\widehat{\sigma} \tag{2.3}$$

where $\widehat{\beta} = T^{-1}\sum_{t=1}^{T}\Delta y_t = (y_T - y_0)/T$ and $\widehat{\sigma}^2 = T^{-1}\sum_{t=1}^{T}(\Delta y_t - \widehat{\beta})^2$.

The test can be generalized by fitting a parametric model and testing the significance of the estimate of the fixed slope β. Either an ARMA model can be fitted to first differences or a structural time series model, based on unobserved component formulation, may be estimated, perhaps by using the STAMP package of Koopman *et al* (2000). The latter may have some attraction when the series is such that it is natural to include components like cycles and seasonals in a model for the levels.

A nonparametric test can be set up in first differences by replacing $\widehat{\sigma}^2$ in (2.3) by an estimate of the long-run variance

$$\hat{\sigma}_L^2(m) = \hat{\gamma}(0) + 2\sum_{\tau=1}^{m} w(\tau, m)\,\hat{\gamma}(\tau) \tag{2.4}$$

where $w(\tau, m)$ is a weighting function, such as the Bartlett window, $w(\tau, m) = 1 - |\tau|/(m+1)$, and $\widehat{\gamma}(\tau)$ is the sample autocovariance of the residuals, $\Delta y_t - \widehat{\beta}$, at lag τ. Alternative options for the kernel $w(.,.)$ and guidelines for choosing m may be found in Andrews (1991). Note that $\widehat{\beta}$ is still asymptotically efficient even though it is computed without taking account of any serial correlation in η_t.

When modified to deal with serial correlation, the distribution of both the parametric and nonparametric $t-$statistics, the latter denoted $t_\beta(m)$, will be asymptotically normal under fairly mild conditions as set out in the first sub-section below. Note that if we were to fit a pure autoregressive model to the differences with lag length determined by some rule of thumb, as is normally done for (2.4), then the test could quite reasonably be classed as nonparametric.

If $\eta_t \sim NID(0, \sigma^2_\eta)$ and $\zeta_t \sim NID(0, \sigma^2_\zeta)$,the LBI test of the null hypothesis of a deterministic slope against the alternative of a nonstationary stochastic slope, that is $H_0 : \sigma^2_\zeta = 0$ against $H_1 : \sigma^2_\zeta > 0$ is to reject for large values of

$$\zeta = T^{-2}\widehat{\sigma}^{-2}\sum_{t=1}^{T}\left(\sum_{i=1}^{t}(\Delta y_i - \widehat{\beta})\right)^2 = T^{-2}\widehat{\sigma}^{-2}\sum_{t=1}^{T}\left(y_t - y_0 - t\widehat{\beta}\right)^2, \tag{2.5}$$

where $\widehat{\sigma}^2$ is as in (2.3); this is the test of Nyblom and Mäkeläinen (1983) applied to first differences. In deriving the test from the LBI principle, one initially obtains the summations running in reverse, that is from $t = T$ to i, but, as a consequence of fitting the slope it can be shown that the two statistics are identical. Asymptotically, ζ has the Cramér-von Mises distribution under the null hypothesis, that is

$$\zeta \xrightarrow{d} \int_0^1 B(r)^2 dr \ , \tag{2.6}$$

where $B(r) = W(r) - rW(1)$, with $W(r)$ being a standard Wiener process or Brownian motion so that $B(r)$ is a Brownian bridge. Gaussianity is not required for the result to hold and the $\eta_t's$ may simply be martingale differences.

If the test statistic is formed without subtracting a deterministic drift it will be locally best invariant (LBI) against $H_1 : \sigma^2_\zeta > 0$ for zero initial conditions, that is $\beta_0 = 0$. However, it is also a consistent test against deterministic drift. Its asymptotic distribution under the null is as in (2.6), but with $B(r)$ replaced by $W(r)$. This distribution is still a member the Cramér-von Mises family, but denoted CvM_0 as opposed to CvM_1 for (2.6). The statistics constructed with forward and reverse partial sums, denoted by ζ^0_F and ζ^0_R respectively, are no longer identical, but have the same asymptotic distribution under the null hypothesis. Although ζ^0_R is the LBI test, it is no more reasonable to treat β_0 as zero than it is to let β_{T+1} be zero, which is why ζ^0_F is considered as well.

To be strictly LBI, the statistics[1] ζ^0_F and ζ^0_R should have the sample mean square rather than the sample variance in the denominator. However, the limiting distribution under the null hypothesis is the same. The local asymptotic

[1]Using the mean square in (2.3) makes it an LBI (LM) test, rather than a Wald test.

distribution is also the same, but because the modified test statistic is bigger[2], it displays higher power in small samples and there appears to be no adverse effect on size.

Parametric and nonparametric forms of the above LBI tests may be constructed. In the nonparametric case, $\widehat{\sigma}_L^2(m)$ replaces $\widehat{\sigma}^2$ just as in the $t-$test; this device is used by Kwiatkowski, Phillips, Schmidt and Shin (1992) in their modification of the test of Nyblom and Mäkeläinen. A parametric statistic can be constructed in the ARIMA framework, as in Leybourne and McCabe (1994), or by fitting a structural time series model, as in Harvey and Streibel (1997).

2.1. Asymptotic distributions under local alternatives

For a model given by (2.1)-(2.2) in which η_t may be serially correlated and/or conditionally heteroscedastic, with positive long run variance, σ_L^2, we consider the limiting behavior of the statistics

$$\begin{aligned} t_\beta(m) &= \sqrt{T}\widehat{\beta}/\widehat{\sigma}(m), \\ \zeta_F^0(m) &= \frac{\sum_{t=1}^T \left(\sum_{i=1}^t \Delta y_i\right)^2}{T^2 \widehat{\sigma}_L^2(m)}, \\ \zeta_R^0(m) &= \frac{\sum_{t=1}^T \left(\sum_{i=t}^T \Delta y_i\right)^2}{T^2 \widehat{\sigma}_L^2(m)}, \\ \zeta(m) &= \frac{\sum_{t=1}^T \left(\sum_{i=1}^t (\Delta y_i - \widehat{\beta})\right)^2}{T^2 \widehat{\sigma}_L^2(m)}, \end{aligned}$$

where $m \to \infty$ such that $m^2/T \to 0$, under the local alternative hypothesis

$$H_{1,T} : \beta_0 = c_d \sigma_L/\sqrt{T},\ \sigma_\zeta^2 = c_s^2 \sigma_L^2/T^2, \tag{2.7}$$

where c_d, c_s are fixed constants. Thus, we consider departures from the null of no drift in the directions of both deterministic and stochastic drift. The following proposition provides the limiting distribution of the statistics under the local alternative hypothesis $H_{1,T}$; a sketched proof is contained in the appendix. These results are then used to evaluate the local asymptotic power of the tests.

[2]When there is only deterministic drift, the probability limit of the denominator of the modified statistic is equal to $(\sigma^2 + \beta^2)/\sigma^2$ times that of the LBI statistic.

Proposition 2.1. *Consider the model (2.1)-(2.2) with η_t being a weakly dependent process as in Stock (1994, p.2745), with long run variance $\sigma_L^2 > 0$. Let $W_0(r)$ and $W_1(r)$ be independent standard Wiener processes for $r \in [0,1]$. Then under $H_{1,T}$,*

$$t_\beta(m) \xrightarrow{d} V(1; c_d, c_s), \tag{2.8}$$

$$\zeta_F^0(m) \xrightarrow{d} \int_0^1 V(r; c_d, c_s)^2 dr, \tag{2.9}$$

$$\zeta_R^0(m) \xrightarrow{d} \int_0^1 \left(V(1; c_d, c_s) - V(r; c_d, c_s)\right)^2 dr, \tag{2.10}$$

$$\zeta(m) \xrightarrow{d} \int_0^1 V^*(r; c_s)^2 dr, \tag{2.11}$$

where

$$\begin{aligned} V(r; c_d, c_s) &= W_0(r) + c_d r + c_s \int_0^r W_1(s) ds, \\ V^*(r; c_s) &= W_0(r) - rW_0(1) + c_s \int_0^r \left(W_1(s) - \int_0^1 W_1(u) du\right) ds. \end{aligned}$$

Remark 1. *For $c_s = 0$, the limiting distribution of $t_\beta^2(m)$ is a noncentral chi-square with one degree of freedom and noncentrality parameter equal to c_d^2, that of $\zeta(m)$ is a standard Cramér-von Mises, CvM_1, as defined in (2.6).*

Remark 2. *For $c_s = c_d = 0$, $t_\beta(m)$ is asymptotically standard normal, while the limiting distributions of $\zeta_F^0(m)$ and $\zeta_R^0(m)$ are CvM_0. The 5% critical value from CvM_0 is 1.656; see table 1 in Nyblom (1989).*

It can also be shown that all tests are consistent against the fixed alternative hypothesis of a stochastic drift, $\sigma_\zeta^2 > 0$, and all but the ζ test are consistent against the alternative hypothesis of deterministic drift. This is confirmed by the computations reported in table 1 below.

2.2. Computation of local asymptotic power

The asymptotic representations given in proposition 2.1 are used to compare the power of the tests against local deviations from the null hypothesis in the direction of deterministic and/or stochastic drift. The results are reported in table 1 in terms of the percentage of rejections. Specifically, we have generated 50000 replications of the limiting random variables defined in (2.8)-(2.11) by replacing the continuous time Wiener processes W_0 and W_1 by their discrete counterparts (dividing the unit interval into 1000 parts) and computing the rejection probabilities for tests run at the 5% level of significance.

As expected, the Wald test, $t_\beta(m)$, is most powerful against a deterministic drift. For example for $c_d = 2$ (and $c_s = 0$), its local asymptotic power is 0.518, as opposed to 0.441 and 0.443 for the tests based on $\zeta_F^0(m)$ and $\zeta_R^0(m)$ respectively. Note that the asymptotic power of the $\zeta(m)$ test against a deterministic drift is always equal to its size.

The $\zeta_R^0(m)$ test achieves the highest power against a stochastic drift starting at zero, that is $c_d = \beta_0 = 0$; indeed it corresponds to the LBI test for this case. Thus with $c_s = 5$, the power of the $\zeta_R^0(m)$ test is 0.569 while that of $t_\beta(m)$ is only 0.524. However, the $t_\beta(m)$ test dominates both the $\zeta_F^0(m)$ and $\zeta(m)$ tests, for which the powers are 0.436 and 0.310 respectively. On the other hand, the power of $\zeta_F^0(m)$ is slightly greater than that $\zeta_R^0(m)$ when c_d is high and c_s is not too large. Of course, if β_{T+1} rather than β_0 had been assumed to be zero, the powers of $\zeta_F^0(m)$ and $\zeta_R^0(m)$ would have been interchanged. Overall, it seems that the Wald test is the best compromise. Even when $c_d = 0$ and $c_s = 50$ its power is only a little below those of the $\zeta_F^0(m)$ and $\zeta_R^0(m)$ tests. Furthermore there may be occasions when a one-sided alternative is plausible, in which case the $t_\beta(m)$ would become even more powerful.

The $\zeta(m)$ test is invariant to c_d and it is dominated by all the other tests except when $c_s = 50$. However, this may be useful insofar as a non-rejection by $\zeta(m)$ and rejection by the other tests is an indication of deterministic drift.

Table 1 Simulated local asymptotic power ($\times$100)

c_s	c_d	0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
0	$t_\beta(m)$	4.9	7.7	16.7	32.2	51.8	70.6	85.2	93.8	97.9
	$\zeta_F^0(m)$	5.0	7.3	14.5	27.2	44.1	62.1	78.0	89.1	95.4
	$\zeta_R^0(m)$	4.9	7.1	14.6	27.4	44.3	62.3	77.8	88.8	95.3
	$\zeta(m)$	4.9	4.9	4.9	4.9	4.9	4.9	4.9	4.9	4.9
1.25	$t_\beta(m)$	11.0	13.9	22.6	35.8	51.7	66.9	80.0	89.4	95.1
	$\zeta_F^0(m)$	8.6	11.1	18.5	30.3	45.4	60.8	75.1	85.9	93.1
	$\zeta_R^0(m)$	12.6	15.0	21.6	32.6	46.1	59.8	72.5	83.1	90.7
	$\zeta(m)$	6.9	6.9	6.9	6.9	6.9	6.9	6.9	6.9	6.9
2.50	$t_\beta(m)$	26.4	28.5	34.1	42.4	52.4	62.5	72.6	81.1	87.7
	$\zeta_F^0(m)$	19.2	21.3	27.4	36.4	47.5	59.1	70.2	79.8	87.2
	$\zeta_R^0(m)$	30.1	31.5	35.5	42.0	49.9	58.0	66.7	74.7	81.8
	$\zeta(m)$	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7	12.7
5.00	$t_\beta(m)$	52.4	53.1	54.8	57.2	60.5	64.2	68.6	73.0	77.3
	$\zeta_F^0(m)$	43.6	44.7	46.9	50.7	55.5	60.8	66.5	72.5	77.9
	$\zeta_R^0(m)$	56.9	57.4	58.4	60.2	62.4	65.0	68.0	71.4	74.8
	$\zeta(m)$	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.0	31.0
10.0	$t_\beta(m)$	74.3	74.2	74.5	75.0	75.6	76.2	76.8	77.8	78.9
	$\zeta_F^0(m)$	70.5	70.7	71.1	71.9	72.9	74.2	75.6	77.4	79.3
	$\zeta_R^0(m)$	78.7	78.9	79.0	79.2	79.6	80.1	80.6	81.3	82.0
	$\zeta(m)$	61.3	61.3	61.3	61.3	61.3	61.3	61.3	61.3	61.3
20.0	$t_\beta(m)$	86.9	87.0	87.0	87.1	87.1	87.2	87.2	87.3	87.4
	$\zeta_F^0(m)$	89.1	89.1	89.3	89.3	89.5	89.8	90.0	90.4	90.6
	$\zeta_R^0(m)$	92.2	92.2	92.4	92.3	92.3	92.5	92.7	92.7	92.8
	$\zeta(m)$	86.9	86.9	86.9	86.9	86.9	86.9	86.9	86.9	86.9
50.0	$t_\beta(m)$	94.8	94.8	94.9	94.9	94.8	94.8	94.8	94.9	94.8
	$\zeta_F^0(m)$	98.8	98.8	98.7	98.8	98.8	98.8	98.8	98.8	98.8
	$\zeta_R^0(m)$	99.1	99.1	99.1	99.1	99.2	99.2	99.1	99.1	99.1
	$\zeta(m)$	99.1	99.1	99.1	99.1	99.1	99.1	99.1	99.1	99.1

3. Permanent drift as the null

A test of the null hypothesis that there is a permanent drift in (2.1) may be based on the simple t_β statistic as defined in (2.3). No modification is needed along the lines of (2.4) since when $\sigma_\zeta^2 > 0$, $t_\beta/\sqrt{T}$ has a limiting distribution that does not depend on the process generating the stationary component, η_t. Specifically,

$$t_\beta/\sqrt{T} \xrightarrow{d} \frac{\int_0^1 W(r)dr}{\left[\int_0^1 \underline{W}(r)^2 dr\right]^{\frac{1}{2}}}, \tag{3.1}$$

where $\underline{W}(r) \equiv W(r) - \int_0^1 W(r)dr$; see sub-section 3.2 below. Note that the numerator is distributed as $N(0, 1/3)$. The test rejects for small values of t_β; that is the critical region is $|t_\beta| < k\sqrt{T}$. The 1%, 5%, 10% critical values, based on the asymptotic distribution, are obtained by setting $k = 0.024$, 0.118, 0.239 respectively. An alternative interpretation of the statistic $t_\beta/\sqrt{T}$ is as the inverse of the coefficient of variation of first differences. We will call this the *standardized drift* and denote it as $\widehat{\beta}^* = \widehat{\beta}/\widehat{\sigma}$. Thus the null hypothesis is rejected at the 5% significance level if $\left|\widehat{\beta}^*\right| < 0.118$.

The above test derives from a proposal made by Bierens (2001) in the context of testing nonstationary cycles. Here the test can be regarded as a test (in differences) at frequency zero. However, we have made a slight modification in that Bierens constructs the denominator without subtracting the mean. In the present context this leads to the statistic, $\widehat{\beta}^\dagger = \widehat{\beta}/(\Sigma\Delta y_t^2/T)^{1/2}$. The asymptotic distribution then has the standard Wiener process, $W(r)$, also in the denominator. This makes virtually no difference to the 5% critical value which is the same as before to three decimal places. Since the statistic is smaller than the one based on $\widehat{\beta}^*$ it will be more likely to reject. However, as the results in the next sub-section show, this appears to make virtually no difference in practice.

What if the drift is purely deterministic? Then

$$\underset{T\to\infty}{p\lim}\widehat{\beta}^* = \beta/\sigma_\eta, \qquad \text{and} \qquad \underset{T\to\infty}{p\lim}\widehat{\beta}^\dagger = \beta/\sqrt{\sigma_\eta^2 + \beta^2}. \tag{3.2}$$

In both cases the null is unlikely to be rejected unless the size of the deterministic drift is small relative to σ_η^2. Specifically, at the 5% level of significance, the null is rejected with probability one as $T \to \infty$ only if $|\beta| < 0.118\sigma_\eta$. It is for this reason that a non-rejection is best regarded as an indication of permanent drift, irrespective of whether it be deterministic or stochastic and nonstationary.

Note that proposition 2.1 immediately shows that the standardized drift test asymptotically has unit power against local deviations from zero drift, as expressed by $H_{1,T}$ of (2.7). For this test the appropriate local asymptotic analysis has to consider local deviations from the null hypothesis of nonstationarity; this is done in sub-section 3.2 below. Before doing that we analyze the finite sample behavior of the test for the model in (2.1)-(2.2).

3.1. Size and power of the standardised drift test for a random walk with stochastic drift.

Consider the random walk with stochastic drift, (2.1)-(2.2), with Gaussian disturbances, and let $q = \sigma_\zeta^2/\sigma_\eta^2$. To evaluate the properties of the standardized drift test, a series of Monte Carlo experiments were carried out, each with 10,000 replications. Table 2 shows the estimated probabilities of rejection for tests at the 5% level of significance, over different values of $q^{1/2}$ for samples of size $T = 50$ and 100. Table 3 includes the corresponding figures for $T = 200$. Results for the augmented Dickey-Fuller (ADF) test with m lags, denoted $\tau(m)$, are given as well; $\tau^*(m)$ indicates the inclusion of a constant. In practice, small values of q are most likely to arise, so the case of $q = 0.01$ ($q^{1/2} = 0.1$) is of particular importance.

The main conclusion to emerge is that the ADF tests tend to reject the null of a nonstationary drift if q is small, i.e. the size is well above the nominal 5%. The reason for this is well-known - the reduced form of second differences contains a moving-average root close to the unit circle and hence the autoregressive approximation is poor. On the other hand the standardized drift test does rather well in that for $T = 100$, the rejection probability is 0.17 for $q = 0.01$ while when $q = 0$, so that the null hypothesis is no longer true, the rejection probability shoots up to 0.76.

These simulations assumed $\beta_0 = 0$. Table 3 illustrates what happens if this is not the case. For non-zero q the rejection probabilities of both the standardized drift tests and $\tau(m)$ are changed very little; $\tau^*(m)$ is unaffected anyway. This is in accordance with the theory of the next sub-section which shows that the local asymptotic distributions are independent of β_0. When $q = 0$, there is a sharp change as β_0 moves from 0.1 to 0.2. This is exactly what one would expect[3] given the probability limit in (3.2).

[3]Note that t_β^2 has a non-central chi-squared distribution with one degree of freedom.

Table 2. Percentage rejections for random walk with stochastic drift, (2.1)

a) T=50

$q^{\frac{1}{2}}$	0	0.1	0.25	0.5	1
$\widehat{\beta}^{\dagger}$	58.5	21.5	10.4	6.8	5.5
$\widehat{\beta}^{*}$	58.2	21.3	10.3	6.7	5.4
$\tau\,(3)$	99.9	88.2	44.5	17.4	7.4
$\tau_{\beta}\,(3)$	97.4	87.0	48.3	19.3	8.6

b) T=100

$q^{\frac{1}{2}}$	0	0.1	0.25	0.5	1
$\widehat{\beta}^{\dagger}$	76.1	16.8	8.0	5.8	5.0
$\widehat{\beta}^{*}$	75.9	16.7	8.0	5.8	4.9
$\tau\,(5)$	100.0	71.6	24.7	8.8	5.3
$\tau_{\beta}\,(5)$	99.7	77.8	26.7	9.4	6.1

Table 3. Percentage rejections for random walk with stochastic drift, (2.1), with non-zero β_0 and $T = 200$

	$q^{\frac{1}{2}}$	0	0.1	0.25	0.5	1
$\beta_0 = 0$	$\widehat{\beta}^{\dagger}$	90.5	12.3	6.8	5.3	5.0
	$\widehat{\beta}^{*}$	90.3	12.2	6.7	5.3	4.9
	$\tau\,(10)$	100.0	36.3	9.5	5.6	5.1
	$\tau^{*}\,(10)$	99.9	41.5	9.6	6.0	6.0
$\beta_0 = 0.1$	$\widehat{\beta}^{\dagger}$	60.3	12.5	6.6	5.3	4.9
	$\widehat{\beta}^{*}$	59.9	12.4	6.6	5.3	4.9
	$\tau\,(10)$	100.0	36.0	9.5	5.6	5.1
	$\tau^{*}\,(10)$	99.9	41.5	9.6	6.0	6.0
$\beta_0 = 0.2$	$\widehat{\beta}^{\dagger}$	12.4	12.7	6.7	5.2	5.0
	$\widehat{\beta}^{*}$	12.1	12.5	6.7	5.2	4.9
	$\tau\,(10)$	99.9	35.6	9.5	5.5	5.2
	$\tau^{*}\,(10)$	99.9	41.5	9.6	6.0	6.0
$\beta_0 = 0.5$	$\widehat{\beta}^{\dagger}$	0.00	10.1	7.0	5.4	4.8
	$\widehat{\beta}^{*}$	0.00	10.0	6.9	5.4	4.7
	$\tau\,(10)$	79.3	31.6	9.3	5.9	5.4
	$\tau^{*}\,(10)$	99.9	41.5	9.6	6.0	6.0

3.2. Local asymptotic power for a local-to-unity autoregressive drift

The standardized drift is a 'pure significance test' in that it is not derived so as to be optimal against any particular alternative. We now consider a difficult situation for it, which is when the drift is stationary but slowly changing. Thus (2.2) is modified to

$$\beta_t = \phi\beta_{t-1} + \zeta_t, \tag{3.3}$$

with β_0 fixed. Local asymptotic power can be analyzed within this framework: the null hypothesis is $H_0^* : \phi = 1$ and the local alternative is $H_{1,T}^* : \phi = 1 - c/T$. Serial correlation in η_t can be handled by incorporating it into the equation for the slope. Thus the disturbance in (3.3) is replaced by

$$\zeta_t^* = \zeta_t + \eta_t - \phi\eta_{t-1}. \tag{3.4}$$

The limiting behavior of the standardized drift under $H_{1,T}^*$ is given by the following proposition; a sketched proof is in the appendix.

Proposition 3.1. *Consider the model (2.1)-(3.3) with ζ_t^* in (3.4) being a weakly dependent process as in Stock (1994, p.2745), with strictly positive long run variance. Then under $H_{1,T}^*$*

$$\widehat{\beta}^* \xrightarrow{d} \frac{\int_0^1 U(r;c)dr}{\left[\int_0^1 \underline{U}(r;c)^2 dr\right]^{\frac{1}{2}}}, \tag{3.5}$$

where $U(r;c)$ is an Ornstein-Uhlenbeck process, defined by the stochastic differential equation

$$dU(r;c) = -cU(r;c)dr + dW(r),$$

with $W(r)$ being a standard Wiener process and $U(0;c) = 0$, that is $U(r;c) = \int_0^r e^{c(s-r)}dW(s)$ as in Phillips (1987), and $\underline{U}(r;c) \equiv U(r;c) - \int_0^1 U(s;c)ds$ is a demeaned Ornstein-Uhlenbeck process.

The limiting distribution of $\widehat{\beta}^\dagger$ is as in (3.5) except that $U(r;c)$ is not demeaned in the denominator.

Remark 3. *On setting $c = 0$, we obtain the limiting null distribution in (3.1). It is also straightforward to show that under the fixed alternative $H_1^* : \phi < 1$, $\widehat{\beta}^* \xrightarrow{p} 0$ so the test is consistent.*

Unit root tests are specifically designed to test the above hypothesis. The local asymptotic distribution of the ADF statistic without a constant included, $\tau(m)$, is

$$\tau(m) \xrightarrow{d} \frac{\int_0^1 U(r;c)dU(r)}{\left[\int_0^1 U(r;c)^2 dr\right]^{\frac{1}{2}}} = \frac{\int_0^1 U(r;c)dW(r)}{\left[\int_0^1 U(r;c)^2 dr\right]^{\frac{1}{2}}} - c\left[\int_0^1 U(r;c)^2 dr\right]^{\frac{1}{2}},$$

where $m \to \infty$ such that $m^3/T \to \infty$. If a constant term is included, the asymptotic representations above holds after replacing $U(r;c)$ by its demeaned version $\underline{U}(r;c)$; see Phillips and Perron (1988, theorem 3) and Stock (1994 p 2772).

The local asymptotic powers of the standardized drift and ADF tests are reported in table 4. As in the previous section, they are computed by simulating the limiting distributions under the local alternative under 50000 replications.

Table 4 Simulated local asymptotic powers of standardized drift and ADF tests for a local-to-unity autoregressive drift.

c	$\widehat{\beta}^*$	$\tau(m)$	$\tau^*(m)$
0	4.9	4.8	4.9
1	6.7	7.7	5.8
2	8.3	11.7	6.6
5	13.4	31.1	11.6
7	16.1	49.1	17.5
10	19.6	75.9	30.7
20	28.4	99.9	85.7
50	44.0	100.0	100.0
100	59.3	100.0	100.0

The $\widehat{\beta}^*$ test has a fairly low probability of rejecting the null hypothesis of a unit root in the drift, though for small values of c, less than ten, it does quite well when set against the ADF test with constant. Its power against moderate serial correlation, c above ten is relatively low.[4] However, the size distortion shown by the ADF tests in tables 2 and 3 means that they are not really a viable alternative.

4. Seasonality

A stochastic seasonal component can be added to a structural time series model, as in Koopman et al (2000), and a t-test carried out on the drift. In a nonparametric framework seasonal differences, rather than first differences, must be taken if the seasonal component is nonstationary. This removes unit roots at the seasonal frequencies. It follows from Busetti and Taylor (2002) that this pre-filtering has no effect on the asymptotic behavior of the tests considered here and it also makes them robust to structural breaks in the seasonal pattern.

If the seasonal component is deterministic, it can be handled by incorporating seasonal dummies into the model for first differences. The asymptotic distribution of ζ and related statistics is unaffected; see Busetti (2002). Thie same is true of the t-tests. Taking seasonal differences will result in a noninvertible process. However, this will not affect any of the tests against permanent drift as the nonparametric correction is based on estimating the spectrum at the origin rather than at the seasonal frequencies. Nor will it affect the asymptotic distribution of the standardized drift test. Hence seasonal differencing may be a safer strategy.

[4]Note that with $c = 100$, setting $T = 100$ gives no serial correlation. The rejection probability of .594 given in table 4 is not inconsistent with the probability of .78 in table 2.

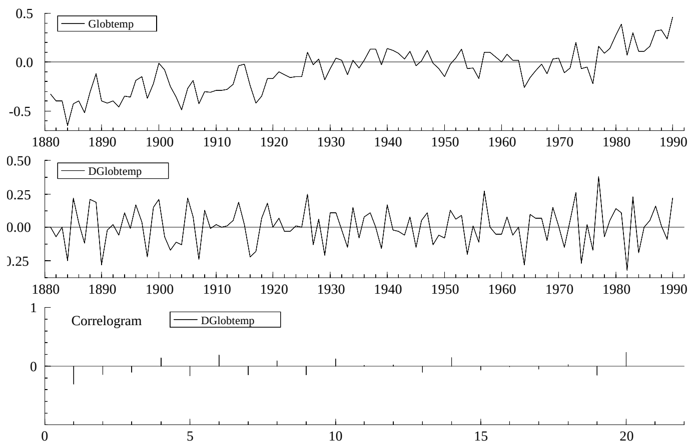

Figure 5.1: Global temperature anomalies: levels, differences and correlogram

5. Examples

Global warming - Figure 1 shows the global annual surface land and marine air anomalies with respect to the 1950-79 average; see Parker *et al* (1995). Ther first difference series is also shown, together with its correlogram.

The nonparametric test statistics for the null of no drift are as follows: $t_\beta(5) = 1.09$; $t_\beta(10) = 1.36$; $\zeta_F^0(5) = .198$; $\zeta_F^0(10) = .308$; $\zeta_R^0(5) = .578$; $\zeta_R^0(10) = 1.055$. None of the tests rejects, though it is interesting that, somewhat unexpectedly, there is a slight tendency for the values to increase as the lag length increases from five to ten. The stochastic drift test tells a consistent story in that $\beta^* = .052$ and $\beta^\dagger = 0.046$, so the null of a permanent drift is rejected.

On the other hand fitting a simple random walk plus drift with an additive irregular component to the levels of the observations[5], using STAMP, gives a $t-$statistic of 1.870. This is close to rejection at the 5% level of significance. A

[5]This is equivalent to modeling the first differences as a first-order moving average. The diagnostics from this model gave no indication of further serial correlation, for example the Box-Ljung statistic based on the first nine residual autocorrelations is Q(9, 8) =8.524. The estimates of the irregular and level variances are 0.0107 and 0.0013, implying a variance for the first differences of 0.0227. The drift is estimated to be 0.0065 which is less than $0.118\sqrt{0.0227}$, hence explaining the rejection with β^*.

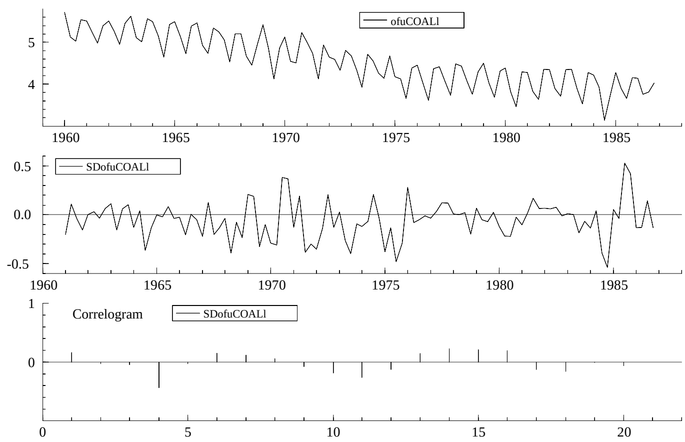

Figure 5.2: Coal consumption of UK Other Final Users

one-sided test would reject and this might be reasonable as it corresponds to a hypothesis of an upward trend in global warming.

Coal- A quarterly series on consumption of coal by 'other final users' in the United Kingdom is used as an example in the manual of the STAMP package of Koopman *et al* (2000). Figure 2 shows the series (in logarithms) together with the seasonally differenced series and its correlogram. It is clear from the graph of the raw series that there is a downward drift which may be best captured by a stochastic slope since the series falls at the outset before levelling off. Fitting a basic structural time series model, consisting of a random walk plus deterministic drift trend, a stochastic seasonal and an irregular yields a $t-$statistic of -3.93 so the hypothesis of no drift is clearly rejected.

The nonparametric test statistics computed from seasonal differences are as follows: $t_\beta(5) = -3.10$; $t_\beta(10) = -3.28$; $\zeta_F^0(5) = 4.40$; $\zeta_F^0(10) = 4.93$; $\zeta_R^0(5) = 3.05$; $\zeta_R^0(10) = 3.42$. The nonparametric $t-$statistic is almost as big as the parametric $t-$statistic and it clearly rejects the null of no slope. Likewise the $\zeta_R^0(m)$ and $\zeta_F^0(m)$ tests reject at the 5% level of significance[6]. The fact that the series appears to have levelled off at the end accounts for the higher values of the forward statistics, i.e. $\beta_{T+1} = 0$ is a better assumption than $\beta_0 = 0$. However, despite this apparent

[6]If the slope is not removed in estimating the variance in the denominator the (LBI) test statistics for $m = 5$ are only 2.85 (forward) and 1.98 (reverse). However, they still reject.

change in the slope, the ζ test is unable to reject the null hypothesis that it is fixed since $\zeta(5) = 0.148$ and $\zeta(10) = 0.166$ and the 5% critical value is 0.461.

The nonparametric test of the null of a permanent drift does not reject as $\beta^* = -0.289$ and $\beta^\dagger = -0.278$ and their absolute values are both greater than the 5% critical value of 0.118. The ADF statistics formed by regressing $\Delta\Delta_4 y_t$ on lagged values and $\Delta_4 y_{t-1}$ are: $\tau(5) = -2.77$, $\tau^*(5) = -3.48$, $\tau(10) = -2.24$, $\tau^*(10) = -3.55$ while the 5% critical values for no constant and constant are -1.96 and -2.89 respectively. Thus all the tests reject the null hypothesis of a unit root in the drift.

6. Conclusion

The Wald t-test is designed to test against a deterministic slope, but it is also consistent against the alternative hypothesis of a stochastic non-stationary slope with an initial value of zero. Overall, it seems to be the best option for testing the null hypothesis of no drift. If it rejects one might apply a standard stationarity test to first differences as this has power only in the direction of a nonstationary drift.

The $t-$statistic can also be used as the basis of a simple nonparametric test of the null hypothesis of nonstationary (permanent) drift. The resulting test is asymptotically less powerful than the ADF test against a local-to-unity autoregressive drift, but in finite samples it does not suffer from the large size distortions of the ADF test when the true underlying process is a random walk plus stochastic drift with a relatively small signal-to-noise ratio.

The examples not only provide illustrations of the tests, but they also point to the type of data generating processes that are plausible. Parametric tests based on fitting structural time series models to the level of a series can often do well as they are designed to capture the effects of irregular and seasonal components added to the level.

Finally, the t-test could also be applied if one wished to test for the presence of a permanent level in a series. This is an unlikely hypothesis for a univariate series. However, it may arise if two series are being compared and the null hypothesis is that they have the same mean. In this case the t-statistic would be computed for the difference between the series. An application in economics is testing whether income per capita in two countries has converged to the same level.

Acknowledgements

The second author thanks the Economic and Social Research Council (ESRC) for support as part of a project on Dynamic Common Factor Models for Regional Time Series, grant number L138 25 1008.

APPENDIX

Proof of Proposition 2.1

The proof follows by standard application of the Functional Central Limit Theorem (FCLT) for dependent sequences and the Continuous Mapping Theorem (CMT). In particular, by the FCLT, under $H_{1,T}$

$$T^{-\frac{1}{2}} \sum_{t=1}^{[Tr]} \Delta y_t \Rightarrow \sigma_L V(r; c_d, c_s), \quad r \in [0,1],$$

where $\Rightarrow$ denotes weak convergence in $D[0,1]$, and the stochastic process $V(r; c_d, c_s)$ is defined in the statement of the proposition. As $\widehat{\sigma}_L^2(m) \xrightarrow{p} \sigma_L^2$, see Stock (1994, page 2799), by an application of the CMT we immediately obtain (2.8) to (2.10). Note that for $c_d = c_s = 0$, $t_\beta(m)$ converges to a standard Normal, $W_0(1)$.

Similar arguments hold for the KPSS-type statistic $\zeta(m)$. Under $H_{1,T}$

$$T^{-\frac{1}{2}} \sum_{t=1}^{[Tr]} \left(\Delta y_t - \widehat{\beta}\right) \Rightarrow \sigma_L V^*(r; c_s),$$

and an application of the CMT delivers (2.11). Note that, since the statistic $\zeta(m)$ is constructed with demeaned first differenced, its limiting distribution is not influenced by the presence of a (local or fixed) drift.

Proof of Proposition 3.1

Under $H_{1,T}^*$ the near integrated process Δy_t weakly converges to an Ornstein-Uhlenbeck process,

$$T^{-\frac{1}{2}} \Delta y_{[Tr]} \Rightarrow \sigma_{\zeta^*,L} U(r; c),$$

where $\sigma^2_{\zeta^*,L}$ is the long run variance of ζ_t^*; see, e.g., Phillips (1987) and Stock (1994, p.2770). Then by application of the CMT

$$
\begin{aligned}
T^{-\frac{1}{2}}\widehat{\beta} &= T^{-\frac{3}{2}}\sum_{t=1}^{T}\Delta y_t \xrightarrow{d} \sigma_{\zeta^*,L}\int_0^1 U(r;c)dr, \\
T^{-1}\widehat{\sigma}^2 &= T^{-2}\sum_{t=1}^{T}(\Delta y_t - \widehat{\beta})^2 \xrightarrow{d} \sigma^2_{\zeta^*,L}\int_0^1 \underline{U}(r;c)^2 dr,
\end{aligned}
$$

which immediately deliver the limiting distribution (3.5) of the standardized drift $\widehat{\beta}^*$. Note the initial value β_0 is asymptotically negligible as it has been assumed to be fixed. The same result would hold true if β_0 were a random variable with bounded second moments; see Tanaka (1996, p. 91).

REFERENCES

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

Busetti, F (2002), Testing for (common) stochastic trends in the presence of structural breaks. *Journal of Forecasting,* 21, 81-105.

Busetti, F. and A.M.R. Taylor (2002), Testing against stochastic trend and seasonality in the presence of unattended breaks and unit roots, University of Birmingham Department of Economics Discussion Paper 02-06.

Bierens, H.J. (2001). Complex unit roots and business cycles: Are they real? *Econometric Theory,* **17,** 962-83.

Harvey, A.C. and M. Streibel (1997), Testing for nonstationary unobserved components, mimeo.

Koopman, S.J., A.C. Harvey, J. A. Doornik and N. Shephard (2000). *STAMP 6.0, Structural Time Series Analyser, Modeller and Predictor.* London: Chapman and Hall.

Kwiatkowski, D., Phillips, P.C.B, Schmidt, P. and Y. Shin (1992), Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root: How Sure Are We that Economic Time Series Have a Unit Root ? *Journal of Econometrics*, **44**, 159-78.

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

Nyblom, J. (1989), Testing for the Constancy of Parameters over Time, *Journal of the American Statistical Association,* **84***, 223-30.*

Nyblom, J. and T. Mäkeläinen (1983), Comparison of tests for the presence of random walk coefficients in a simple linear model, *Journal of the American Statistical Association,* **78***, 856-64.*

Parker, D.E., Folland, C.K. and M. Jackson (1995). Marine surface temperature: observed variations and data requirements. *Climatic Change* 31, 559-600.

Phillips, P.C.B. (1987), Time series regression with a unit root, *Econometrica* 55, 277-301.

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

Stock, J.H. (1994) Unit roots, structural breaks and trends, in R.F. Engle and D.L. McFadden (eds.), *Handbook of Econometrics* 4, Elsevier Science, 2739-2840.

Tanaka, K., 1996, *Time series analysis: nonstationary and noninvertible distribution theory*, Wiley: New York.