January 26, 2001

Unit Root Tests for Time Series with Level Shifts: A Comparison of Different Proposals

by

Markku Lanne
University of Helsinki

and

Helmut Lütkepohl
Humboldt University Berlin

Abstract

A number of unit root tests which accommodate a deterministic level shift at a known point in time are compared in a Monte Carlo study. The tests differ in the way they treat the deterministic term of the DGP. It turns out that Phillips-Perron type tests have very poor small sample properties and cannot be recommended for applied work. Moreover, tests which estimate the deterministic term by a GLS procedure under the unit root null hypothesis are superior in terms of size and power properties relative to tests which estimate the deterministic term by OLS procedures.

JEL classification: C22, C12

Keywords: Univariate time series, unit root, structural shift, autoregression

*Comments by Luis Gil-Alaña are gratefully acknowledged. We thank the Deutsche Forschungsgemeinschaft, SFB 373, and the European Commission under the Training and Mobility of Researchers Programme (contract No. ERBFMRXCT980213) for financial support. Part of this research was carried out while the first author was visiting the Humboldt University in Berlin.

1 Introduction

Testing for unit roots is common practice at the beginning of any analysis of economic time series because the trending properties of the series determine to some extent the models and inference procedures to be used in later stages of the analysis. Suitable unit root tests for different situations have been proposed. One group of tests allows for a level shift in the data generation process (DGP). Surprisingly little seems to be known about the relative performance of some of the tests in this class. In this study we intend to compare some of these tests and thereby we will close this gap at least partly.

There are a number of tests which allow for various kinds of shifts in the mean of the DGP (see, e.g., Lanne, Lütkepohl & Saikkonen (1999) for a range of tests allowing for very general shift functions). However, the original proposals focus on the case where there is a one time shift which can be captured by a dummy variable (e.g., Perron (1989, 1990)). Moreover, procedures are available for cases where the time of the shift is unknown. In this study we compare unit root tests for DGPs with a shift in mean which can be described by a shift dummy and for which the break date is known. The reason is that we expect tests which perform well under these relatively simple conditions are also preferable in more complex situations. In turn, tests which do not perform well in simple situations are not expected to do well in difficult circumstances. Assuming a known shift date is also of practical relevance because there are many time series with changes in their construction or definition which occur at a known point in time. Examples are German macroeconomic time series which have a shift at the time of the German unification or exchange rates for which currency adjustments have occurred.

In the next section we present the tests without explicitly describing the DGP. Of course, a special DGP is assumed in deriving the tests. In practice, however, for a given time series a special test is usually applied under the assumption that it is suitable for the time series at hand. It is a different question whether the true DGP has actually generated the given time series. Clearly, the true DGP will be unknown because otherwise a test would be unnecessary.

The study is organized as follows. A number of unit root tests allowing for a level shift are reviewed in the next section. In Section 3 a Monte Carlo comparison is performed and conclusions follow in Section 4.

2 The Tests

The tests are described for the case where a linear time trend is included in the model. If a time trend is not required for a time series of interest, it can simply be dropped. The adjustments are obvious and straightforward. Different tables for the critical values have to be used in that case, however. It is also assumed that the shift can be modeled by a dummy variable

$$DU_t = \begin{cases} 0, & t \leq T_B \\ 1, & t > T_B \end{cases},$$

where T_B is the shift date.

In the time series literature two generating mechanisms of shifts are distinguished, additive outlier models and innovational outlier models. The former result in an abrupt shift in the level, whereas the latter allow for a smooth shift from the initial level to a new level. We will treat tests for the two model types in turn starting with additive outlier models.

Additive Outlier Assumption

Most of the tests derived in the additive outlier framework may be viewed as adjusting the original time series, say $y_1, \ldots, y_T$, in a first step. Then augmented Dickey-Fuller (ADF) or Phillips-Perron (PP) tests are applied to the adjusted series. In other words, denoting the adjusted series by $\tilde{y}_1, \ldots, \tilde{y}_T$, the following four test versions are considered.

- **ADF tests**

  Consider a regression

  $$\Delta\tilde{y}_t = \phi\tilde{y}_{t-1} + \sum_{j=1}^{p} \eta_j \Delta DU_{t-j} + \sum_{j=1}^{p} b_j \Delta\tilde{y}_{t-j} + error_t \tag{2.1}$$

  and denote the OLS estimator of ϕ and its t-statistic by $\tilde{\phi}$ and $t_{\tilde{\phi}}$, respectively. Using the notation $\tilde{b}(1) = 1 - \tilde{b}_1 - \cdots - \tilde{b}_p$, where $\tilde{b}_j$ is the OLS estimator of b_j $(j = 1, \ldots, p)$, the ADF statistics are $T\tilde{\phi}/\tilde{b}(1)$ and $t_{\tilde{\phi}}$ (see Perron & Vogelsang (1993)).

- **PP tests**

  Denote the residuals of a regression $\tilde{y}_t = \alpha\tilde{y}_{t-1} + e_t$ by $\tilde{e}_t$ $(t = 1, \ldots, T)$ and define

  $$\tilde{\sigma}_e^2 = T^{-1}\sum_{t=1}^{T} \tilde{e}_t^2$$

and

$$\tilde{\sigma}^2 = T^{-1}\sum_{t=1}^{T}\tilde{e}_t^2 + 2T^{-1}\sum_{j=1}^{k}\omega_k(j)\sum_{t=j+1}^{T}\tilde{e}_t\tilde{e}_{t-j}$$

with $\omega_k(j) = 1 - \frac{j}{k+1}$ (Newey-West estimator). Then two versions of PP tests are

$$Z(\tilde{\alpha}) = T(\tilde{\alpha}-1) - \frac{1}{2}(\tilde{\sigma}^2 - \tilde{\sigma}_e^2)\left(T^{-2}\sum_{t=2}^{T}\tilde{y}_{t-1}^2\right)^{-1}$$

and

$$Z(t_{\tilde{\alpha}}) = \frac{\tilde{\sigma}_e}{\tilde{\sigma}}t_{\tilde{\alpha}} - \frac{1}{2}(\tilde{\sigma}^2 - \tilde{\sigma}_e^2)\left[\tilde{\sigma}\left(T^{-2}\sum_{t=2}^{T}\tilde{y}_{t-1}^2\right)^{1/2}\right]^{-1}$$

(see Phillips & Perron (1988) for the tests and Perron & Vogelsang (1993) for critical values).

The following possibilities for adjusting y_t have been proposed in the literature. Perron (1989, 1990) proposes to run a regression

$$y_t = \mu_0 + \mu_1 t + \delta DU_t + error_t \quad (t = 1, \ldots, T). \tag{2.2}$$

The regression residuals $\tilde{y}_t^P = y_t - \tilde{\mu}_0 - \tilde{\mu}_1 t - \tilde{\delta}DU_t$ $(t = 1, \ldots, T)$ are the adjusted y_t. The ADF versions $T\tilde{\phi}/\tilde{b}(1)$ and $t_{\tilde{\phi}}$ of the resulting unit root test statistics will be denoted by P_{AO}^{ϕ} and P_{AO}^{t}, respectively, and the $Z(\tilde{\alpha})$ and $Z(t_{\tilde{\alpha}})$ versions are denoted as PP_{AO}^{α} and PP_{AO}^{t}, respectively.

Based on a unit root test of Schmidt & Phillips (1992), Amsler & Lee (1995) use a Lagrange Multiplier (LM) approach and propose to estimate μ_1 and δ under the null hypothesis, that is, (2.2) is multiplied by the differencing operator Δ and then an OLS regression

$$\Delta y_t = \mu_1 + \delta\Delta DU_t + error_t \quad (t = 2, \ldots, T) \tag{2.3}$$

is performed. Denoting the resulting parameter estimators by $\tilde{\mu}_1$ and $\tilde{\delta}$ and defining $\tilde{\mu}_x = y_1 - \tilde{\mu}_1$, the adjusted series is $\tilde{y}_t^{AL} = y_t - \tilde{\mu}_x - \tilde{\mu}_1 t - \tilde{\delta}DU_t$. The resulting test statistics corresponding to $T\tilde{\phi}/\tilde{b}(1)$, $t_{\tilde{\phi}}$, $Z(\tilde{\alpha})$ and $Z(t_{\tilde{\alpha}})$ will be denoted as AL_{ADF}^{ϕ}, AL_{ADF}^{t}, AL_{PP}^{α} and AL_{PP}^{t}, respectively.

Extending the ideas of Saikkonen & Lütkepohl (1999), Lanne, Lütkepohl & Saikkonen (1999) (henceforth LLS) propose to estimate the parameters of model (2.2) by a GLS approach under local alternatives or the unit root null hypothesis as in (2.3). We focus on estimation under the null hypothesis because in small sample simulations this turned out to work better than estimation under local alternatives (see LLS). Assuming that the errors in (2.3) have an AR(p) serial correlation structure, they propose to estimate the relevant parameters $\theta = (\mu_0, \mu_1, \delta)$ by minimizing

$$Q_p(\theta, b) = (Y - Z\theta)'\Sigma_p(b)^{-1}(Y - Z\theta), \tag{2.4}$$

where $\Sigma_p(b) = \sigma^{-2}\mathrm{Cov}(U)$ with U being the error vector of (2.4), $Y = [y_1 : (y_2 - y_1) : \cdots : (y_T - y_{T-1})]'$ and $Z = [Z_1 : Z_2 : Z_3]$ with $Z_1 = [1, 0, \ldots, 0]'$, $Z_2 = [1, 1, \ldots, 1]'$ and $Z_3 = [DU_1 : (DU_2 - DU_1) : \cdots : (DU_T - DU_{T-1})]'$. Although the adjusted data $\tilde{y}_t^{LLS} = y_t - \tilde{\mu}_0 - \tilde{\mu}_1 t - \tilde{\delta} DU_t$ could be used in the ADF and PP approaches, LLS propose a slightly different procedure which adjusts for the estimation errors in the nuisance parameters and worked quite well in small sample simulations. Denoting the AR polynomial by $b(L) = 1 - b_1 L - \cdots - b_p L^p$ and its estimator from (2.4) by $\tilde{b}(L)$, LLS define $\tilde{w}_t = \tilde{b}(L)\tilde{y}_t^{LLS}$ and base the unit root test on the auxiliary regression model

$$\Delta\tilde{w}_t = \nu + \phi\tilde{w}_{t-1} + \pi\tilde{b}(L)\Delta DU_t + \sum_{j=1}^{p} \alpha_j \Delta\tilde{y}_{t-j}^{LLS} + error_t \quad (t = p+2, \ldots, T). \tag{2.5}$$

The unit root test statistic is again obtained as the usual t-statistic of the estimator of ϕ based on OLS estimation of this model. It will be denoted by LLS_{AO}.

Innovational Outlier Assumption

Under the assumption of an innovational outlier the parameters of the deterministic terms can be estimated jointly with the parameter on which the unit root test is based. This was proposed by Perron (1989). He suggests using the regression model

$$\Delta y_t = \phi y_{t-1} + \mu_0 + \mu_1 t + \delta DU_t + \delta^* \Delta DU_t + \sum_{j=1}^{p} b_j \Delta y_{t-j} + error_t \quad (t = p+2, \ldots, T) \tag{2.6}$$

without prior adjustment of the data series. Again the t-statistic of ϕ is the relevant unit root statistic. As mentioned earlier, the advantage of model (2.6) is that the level shift is gradual rather than abrupt. The resulting ADF type unit root test statistics will be denoted by P_{IO}^{ϕ} and P_{IO}^{t}.

Based on ideas of Lütkepohl, Müller & Saikkonen (1999), LLS also consider a second possibility for estimating the nuisance parameters under the unit root null hypothesis using a regression model

$$\Delta y_t = \mu_0 D_t + \mu_1 + \delta \Delta DU_t + \sum_{j=1}^{p} b_j \Delta y_{t-j} + error_t \quad (t = 1, \ldots, T), \tag{2.7}$$

where $D_t = 1$ for $t = 1$ and zero elsewhere and $y_t = 0$ for $t < 1$. From this model the parameters are estimated by OLS. The unit root test is based on $\tilde{v}_t = \tilde{b}(L)y_t - \tilde{\mu}_0 - \tilde{\mu}_1 t - \tilde{\delta} DU_t$ using the auxiliary regression

$$\Delta \tilde{v}_t = \nu + \phi \tilde{v}_{t-1} + \Delta DU_t \pi_1 + \tilde{q}_t' \pi_2 + error_t \quad (t = 2, \ldots, T), \tag{2.8}$$

where $\tilde{q}_t' = [\Delta y_{t-1} - \tilde{\mu}_* : \cdots : \Delta y_{t-p+1} - \tilde{\mu}_*]$ and $\tilde{\mu}_* = \tilde{\mu}_1 / \tilde{b}(1)$. Again the t-statistic of the estimated ϕ is used as the unit root test statistic. It will be denoted by LLS_{IO}. All the tests are listed in Table 1. In the next section we will explore their small sample properties using a Monte Carlo experiment.

Table 1. Unit Root Tests

Test statistic	Estimation of deterministic terms	References
P_{AO}^{ϕ}	OLS	Perron (1989, 1990)
P_{AO}^{t}	OLS	Perron & Vogelsang (1993)
PP_{AO}^{α}	OLS	
PP_{AO}^{t}	OLS	
AL_{ADF}^{ϕ}	OLS under H_0	Amsler & Lee (1995)
AL_{ADF}^{t}	OLS under H_0	Schmidt & Phillips (1992)
AL_{PP}^{α}	OLS under H_0	
AL_{PP}^{t}	OLS under H_0	
LLS_{AO}	GLS under H_0	Lanne, Lütkepohl & Saikkonen (1999)
P_{IO}^{ϕ}	OLS of full model	Perron (1989, 1990)
P_{IO}^{t}	OLS of full model	
LLS_{IO}	GLS under H_0 in full model	Lanne, Lütkepohl & Saikkonen (1999)

3 Monte Carlo Comparison

We have compared the unit root tests using the following two DGPs:

$$y_t = DU_t + x_t, \qquad (1 - b_1 L)(1 - \rho L)x_t = \varepsilon_t \qquad (t = 1, \ldots, T) \tag{3.1}$$

and

$$(1 - b_1 L)y_t = DU_t + v_t, \qquad v_t = \rho v_{t-1} + \varepsilon_t \qquad (t = 1, \ldots, T) \tag{3.2}$$

with $\varepsilon_t \sim iid\, N(0,1)$, $\rho = 1, 0.9, 0.8$ and $T = 100, 200$. Processes of this type were also used in a Monte Carlo study by LLS. In addition to the sample values we have generated 100 presample values which were discarded except that some of them are used in the estimation procedures for which presample values are required. The process (3.1) represents an additive outlier model with an abrupt shift at time T_B whereas the model (3.2) corresponds to an innovational outlier process if $b_1 \neq 0$. We apply all tests to the time series generated by the two different DGPs although, strictly speaking, the tests designed for additive outliers are not constructed to perform well for time series with innovational outliers and vice versa. In practice, however, the type of DGP will be unknown in general and, hence, it would be advantageous for a test to be robust with respect to the DGP. This is one issue to be addressed in the simulation study.

Although the DGPs do not have deterministic linear trend terms, we have applied the tests which allow for such a term and present some results for sample size $T = 100$, break point $T_B = 49$ and different values of b_1 in Tables 2 and 3. The figures in the tables are rejection frequencies for tests with nominal significance level of 5% based on 1000 replications of the experiment. Thus, the Monte Carlo standard error is $\sqrt{P(1-P)/1000}$ for a true rejection probability P. For instance, for $P = 0.05$ we get a standard error of 0.007.

A striking observation from Tables 2 and 3 is the very poor performance of the PP tests. They are extremely conservative and, as a result, have very little power for $b_1 > 0$, whereas they overreject dramatically under H_0 for $b_1 < 0$ (see the $\rho = 1$ columns). Size problems of PP tests were also observed by Schwert (1989) for DGPs without structural shifts. Although the finite sample properties of these tests improve slightly for $T = 200$, their actual rejection frequencies in the presence of a unit root are still far from the nominal 5% even for such relatively large samples. In Tables 2 and 3, a truncation lag of $k = 2$ has been used for the tests. The problems persist, however, for other truncation lags such as $k = 4, 6, 8$. Clearly, these findings make the tests useless for applied work.

The P^{ϕ}_{AO} test also has size problems. It overrejects a bit too much in some situations (see, e.g., $b_1 = 0.8$) while its power is often no better than that of LLS_{AO} and LLS_{IO}. Notice that the power figures are not size-adjusted. Hence, P^{ϕ}_{AO} is clearly inferior to LLS_{AO} and LLS_{IO} and, consequently, it cannot be recommended for empirical work either.

Since all other tests are conservative in some situations, it may not be surprising that they are also inferior in terms of power to LLS_{AO} and LLS_{IO} which tend to overreject slightly in some cases. The magnitude of the gains in power obtained by using the latter tests is remarkable, however. It is also noteworthy that the tests are superior even if they are applied in situations where their underlying assumptions regarding the DGP are violated. In other words, LLS_{AO} performs relatively well for DGP (3.2) and the same is true for LLS_{IO} with respect to the DGP (3.1) although these processes are not the ones assumed in the derivation of the tests, respectively. Thus, overall LLS_{AO} and LLS_{IO} are clearly the tests with the best performance in our simulations.

4 Conclusions

We have compared a range of unit root tests allowing for a structural break in their mean. The tests are based on different methods for estimating the deterministic terms including the level shift. They also differ in the way they account for short-term dynamics. Specifically, ADF type tests which take care of short-term dynamics in a parametric way and Phillips-Perron versions which use nonparametric adjustments for short-term dynamics are considered. It turns out that Phillips-Perron type tests perform very poorly under our simulation setup. They are very conservative and consequently also have very low power in some situations, whereas in other cases they overreject dramatically. Among the other tests, those which estimate the deterministic part of the DGP by a GLS procedure under the unit root null hypothesis have by far the best small sample properties overall. They have roughly correct size although they sometimes overreject slightly. Moreover, their power is far better than that of those competitors which also respect the nominal significance level.

References

Amsler, C. & J. Lee (1995), An LM test for a unit root in the presence of a structural change, *Econometric Theory*, 11, 359 - 368.

Lanne, M., H. Lütkepohl & P. Saikkonen (1999), Comparison of unit root tests for time series with level shifts, Discussion Paper 88, SFB 373, Humboldt-Universität zu Berlin.

Lütkepohl, H., C. Müller & P. Saikkonen (1999), Unit root tests for time series with a structural break when the break point is known, Discussion Paper, SFB 373, Humboldt-Universität zu Berlin.

Perron, P. (1989), The great crash, the oil price shock and the unit root hypothesis, *Econometrica*, 57, 1361 - 1401.

Perron, P. (1990), Testing for a unit root in a time series with a changing mean, *Journal of Business & Economic Statistics*, 8, 153 - 162.

Perron, P. & T.J. Vogelsang (1993), Erratum, *Econometrica*, 61, 248 - 249.

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

Saikkonen, P. & H. Lütkepohl (1999), Testing for unit roots in time series with level shifts, Discussion Paper, SFB 373, Humboldt-Universität zu Berlin.

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

Schwert, G.W. (1989), Tests for unit roots: A Monte Carlo investigation, *Journal of Business & Economic Statistics*, 7, 147 - 159.

Table 2. Frequency Distributions of Unit Root Tests with Linear Trend Based on 1000 Monte Carlo Replications, Sample Size $T = 100$ and Break Point $T_B = 49$

Test	DGP (3.1)			DGP (3.2)		
	$\rho = 1.0$	$\rho = 0.9$	$\rho = 0.8$	$\rho = 1.0$	$\rho = 0.9$	$\rho = 0.8$
	$b_1 = 0.5$					
P^{ϕ}_{AO}	0.094	0.205	0.407	0.094	0.196	0.405
P^{t}_{AO}	0.051	0.137	0.321	0.051	0.137	0.316
PP^{α}_{AO}	0.001	0.001	0.014	0.001	0.002	0.014
PP^{t}_{AO}	0.003	0.007	0.028	0.003	0.009	0.028
AL^{ϕ}_{ADF}	0.021	0.116	0.318	0.021	0.111	0.314
AL^{t}_{ADF}	0.017	0.093	0.267	0.019	0.085	0.258
AL^{α}_{PP}	0.001	0.005	0.057	0.000	0.004	0.054
AL^{t}_{PP}	0.000	0.005	0.046	0.000	0.003	0.043
LLS_{AO}	0.080	0.233	0.526	0.075	0.216	0.499
P^{ϕ}_{IO}	0.031	0.138	0.337	0.032	0.137	0.338
P^{t}_{IO}	0.053	0.131	0.310	0.056	0.139	0.316
LLS_{IO}	0.081	0.217	0.455	0.079	0.216	0.468
	$b_1 = -0.5$					
P^{ϕ}_{AO}	0.045	0.167	0.560	0.045	0.168	0.564
P^{t}_{AO}	0.059	0.150	0.508	0.058	0.153	0.503
PP^{α}_{AO}	0.387	0.775	0.990	0.384	0.779	0.990
PP^{t}_{AO}	0.495	0.865	0.996	0.501	0.867	0.996
AL^{ϕ}_{ADF}	0.024	0.094	0.335	0.023	0.093	0.336
AL^{t}_{ADF}	0.018	0.072	0.276	0.018	0.078	0.280
AL^{α}_{PP}	0.180	0.505	0.795	0.175	0.502	0.799
AL^{t}_{PP}	0.199	0.546	0.824	0.196	0.557	0.822
LLS_{AO}	0.078	0.284	0.680	0.079	0.284	0.699
P^{ϕ}_{IO}	0.018	0.110	0.449	0.018	0.109	0.435
P^{t}_{IO}	0.061	0.158	0.484	0.057	0.146	0.466
LLS_{IO}	0.070	0.228	0.578	0.075	0.250	0.610

Table 3. Frequency Distributions of Unit Root Tests with Linear Trend Based on 1000 Monte Carlo Replications, Sample Size $T = 100$ and Break Point $T_B = 49$

Test	DGP (3.1)			DGP (3.2)		
	$\rho = 1.0$	$\rho = 0.9$	$\rho = 0.8$	$\rho = 1.0$	$\rho = 0.9$	$\rho = 0.8$
	$b_1 = 0.8$					
P_{AO}^{ϕ}	0.161	0.227	0.269	0.158	0.218	0.266
P_{AO}^{t}	0.032	0.077	0.170	0.030	0.074	0.147
PP_{AO}^{α}	0.000	0.000	0.000	0.000	0.000	0.001
PP_{AO}^{t}	0.001	0.005	0.001	0.001	0.005	0.004
AL_{ADF}^{ϕ}	0.024	0.085	0.170	0.025	0.079	0.145
AL_{ADF}^{t}	0.015	0.071	0.130	0.018	0.066	0.121
AL_{PP}^{α}	0.000	0.000	0.001	0.000	0.000	0.000
AL_{PP}^{t}	0.000	0.000	0.001	0.000	0.000	0.000
LLS_{AO}	0.065	0.167	0.286	0.063	0.149	0.262
P_{IO}^{ϕ}	0.035	0.097	0.178	0.040	0.096	0.185
P_{IO}^{t}	0.056	0.116	0.181	0.054	0.112	0.184
LLS_{IO}	0.077	0.159	0.268	0.079	0.161	0.269
	$b_1 = -0.8$					
P_{AO}^{ϕ}	0.045	0.164	0.583	0.046	0.167	0.588
P_{AO}^{t}	0.052	0.153	0.517	0.050	0.151	0.512
PP_{AO}^{α}	0.883	0.998	1.000	0.887	0.998	1.000
PP_{AO}^{t}	0.927	0.999	1.000	0.926	0.999	1.000
AL_{ADF}^{ϕ}	0.022	0.061	0.210	0.021	0.068	0.212
AL_{ADF}^{t}	0.020	0.050	0.180	0.018	0.056	0.175
AL_{PP}^{α}	0.584	0.880	0.972	0.588	0.885	0.969
AL_{PP}^{t}	0.627	0.907	0.981	0.627	0.905	0.980
LLS_{AO}	0.081	0.294	0.696	0.080	0.292	0.719
P_{IO}^{ϕ}	0.066	0.238	0.696	0.060	0.230	0.686
P_{IO}^{t}	0.060	0.162	0.526	0.051	0.150	0.500
LLS_{IO}	0.069	0.227	0.579	0.073	0.250	0.614