Economics Division
University of Southampton
Southampton SO17 1BJ, UK

Discussion Papers in
Economics and Econometrics

Title Jointly Testing Linearity and Nonstationarity within Threshold Autoregressions

By Jean-Yves Pitarakis (University of Southampton)

No. 1206

ISSN 0966-4246

Jointly Testing Linearity and Nonstationarity within Threshold Autoregressions*

Jean-Yves Pitarakis
University of Southampton
Economics Division
United-Kingdom

May 1, 2012

Abstract

We develop a test of the joint null hypothesis of linearity and nonstationarity within a threshold autoregressive process of order one with deterministic components. We derive the limiting distribution of a Wald type test statistic and subsequently investigate its local power and finite sample properties. We view our test as a useful diagnostic tool since a non rejection of our null hypothesis would remove the need to explore nonlinearities any further and support a linear autoregression with a unit root.

Keywords: Threshold Autoregressive Models, Unit Roots, Near Unit Roots, Brownian Bridge, Augmented Dickey Fuller Test.

JEL: C2, C5, C12.

*Address for Correspondence: University of Southampton, School of Social Sciences, Economics Division, Southampton SO17 1BJ, United-Kingdom. Email: J.Pitarakis@soton.ac.uk. Tel: +44-23-80592631. Financial Support from the ESRC through the research grant RES-000-22-3983 is gratefully acknowledged.

1 Introduction

This paper is concerned with inferences within an environment that combines threshold type nonlinearities with the presence of a highly persistent variable that contains a unit root. One of the first papers to introduce an environment that combined unit root type of nonstationarities with nonlinear dynamics was Caner and Hansen (2001). Operating within an autoregressive specification formulated as an ADF regression the authors developed two key tests for detecting the presence of threshold effects when the underlying variable contains a unit root under the null hypothesis (see also Pitarakis (2008)). Their first test was designed to test the null of linearity in *all* the parameters of the ADF regression without explicitly imposing the unit root restriction within the null hypothesis. A random walk with drift was however maintained as the data generating process. In a second test the authors concentrated solely on the autoregressive parameters associated with the presence or absence of a unit root and developed tests of the joint null of a unit root and linearity without constraining the remaining parameters of the ADF regression associated with the deterministic components (i.e. constant and trend).

In this paper we argue that a useful addition to the existing toolkit for uncovering threshold effects in nonstationary environments is a test that would allow one to test the *joint* null of linearity in all the parameters of the ADF regression *and* nonstationarity. In this context we are interested in the limiting distribution of a Wald type test under a null hypothesis that imposes not only the stability of all AR parameters but also the unit root explicitly. We expect such a test to have power against departures from linearity as well as departures from the unit root null. More importantly a non rejection of this joint null would conclude the analysis and support the modelling of the variable under investigation through a linear unit root process. In this sense it may be viewed as a useful diagnostic tool before attempting to undertake any further investigation of nonlinear dynamics.

2 The Model and Asymptotic Inference

We are interested in testing $H_0^A : \theta_1 = \theta_2, \rho_1 = \rho_2 = 0$ in

$$\Delta y_t \quad = \quad (\theta_1' w_{t-1} + \rho_1 y_{t-1}) I(Z_{t-1} \leq \gamma) + (\theta_2' w_{t-1} + \rho_2 y_{t-1}) I(Z_{t-1} > \gamma) + e_t \tag{1}$$

with $w_{t-1} = (1\ t)'$ and $\theta_i = (\mu_i\ \delta_i)'$ for $i = 1, 2$. $Z_t = y_t - y_{t-m}$ with $m \geq 1$ is the stationary threshold variable and the threshold parameter γ is assumed unknown with $\gamma \in \Gamma = [\gamma_1, \gamma_2]$. The parameters γ_1 and γ_2 are selected such that $P(Z_t \leq \gamma_1) = \pi_1 > 0$ and $P(Z_t \leq \gamma_2) = \pi_2 < 1$ for given π_1 and π_2 (e.g. 10% trimming on both ends). As in Caner and Hansen (2001) and for later use it is also convenient to rewrite $I(Z_{t-1} \leq \gamma) = I(G(Z_{t-1}) \leq G(\gamma)) \equiv I(U_{t-1} \leq \lambda)$ where $G(.)$ is the marginal distribution of Z_t and U_t denotes a uniformly distributed random variable on $[0, 1]$. Throughout this paper and for notational simplicity we also let I_{1t-1} and I_{2t-1} denote the two indicator functions $I(U_{t-1} \leq \lambda)$ and $I(U_{t-1} > \lambda)$.

Letting $\Psi_i = (\mu_i\ \delta_i\ \rho_i)'$, in Caner and Hansen (2001) the authors derived the limiting behaviour of a Wald type test statistic for testing $H_0 : \Psi_1 = \Psi_2$ in (1) when the underlying process is known to contain an exact unit root with or without an intercept (e.g. $\Delta y_t = \mu + e_t$). Proceeding under the same probabilitic assumptions our goal here is to instead develop inferences for testing the joint null hypothesis of linearity and unit root $H_0^A : \theta_1 = \theta_2, \rho_1 = \rho_2 = 0$ via a Wald type test statistic.

We rewrite (1) in matrix form as $\Delta Y = X_1\Psi_1 + X_2\Psi_2 + e$ with X_i stacking the elements given by I_{it-1}, tI_{it-1}, $y_{t-1}I_{it-1}$. Letting $V = [X_1\ X_2]$ we also write $\Delta Y = V\Psi + e$ with $\Psi = (\Psi_1\ \Psi_2)'$ so that the Wald statistic associated with $H_0^A : \mu_1 = \mu_2, \delta_1 = \delta_2, \rho_1 = \rho_2 = 0$ is $W_T^A(\lambda) = \hat{\Psi}' R_A' [R_A (V'V)^{-1} R_A']^{-1} R_A \hat{\Psi}/\hat{\sigma}^2$ with $R_A = \{(1,0,0,-1,0,0), (0,1,0,0,-1,0), (0,0,1,0,0,0), (0,0,0,0,0,1)\}$. Here $\hat{\sigma}^2$ refers to the unrestricted residual variance. Before stating our main results we also let $DF_{\tau,\infty}$ denote the limiting distribution of the t-ratio for testing $H_0 : \rho = 0$ in $\Delta y_t = \mu + \delta t + \rho y_{t-1} + e_t$ as stated in Hamilton (1988, pp. 549-550, Equations (17.4.53) and (17.4.54)). See also Phillips and Perron (1888, Theorem 1(e) with $\lambda = 0$ and $\sigma/\sigma_u = 1$). The limiting behaviour of the supremum version of $W_T^A(\lambda)$ is now summarised in the following Proposition with the supremum understood to be taken over some symmetric interval $\Lambda = [\lambda_0, 1 - \lambda_0]$.

Proposition 1. *Under the same assumptions as in Caner and Hansen (2001) and under* $H_0^A : \theta_1 = \theta_2, \rho_1 = \rho_2 = 0$ *we have as* $T \to \infty$,

$$\sup_{\lambda} W_T^A(\lambda) \quad \Rightarrow \quad \sup_{\lambda} BB(\lambda)/\lambda(1-\lambda) + DF_{\tau,\infty}^2 \tag{2}$$

with $BB(\lambda)$ *denoting a standard Brownian Bridge process of the same dimension as* ϕ_i.

The above limit is free of nuisance parameters. Its first component is the familiar normalised squared Brownian Bridge type of process while the second one arises due to the unit root imposed within the null hypothesis. More specifically $DF_{\tau,\infty} = [\int_0^1 BdB + A]/\sqrt{D}$ with A and D as in Phillips and Perron (1988) and B the standard Brownian Motion associated with the $e_t's$.

We expect that the above test will have nontrivial power against departures from linearity as well as the unit root null. At this stage it is interesting to contrast the above limit with the one that occurs within a similar setting but with structural break based regimes instead of thresholds in (1). In Pitarakis (2011) the author has investigated a similar null hypothesis within an ADF regression with a structural break and documented a limiting distribution composed also of two components one of which was again given by $DF_{\tau,\infty}^2$ but with its first component being nonstandard and substantially different from the Brownian Bridge limit above. This highlights the fundamentally different asymptotics that results from alternative approaches of capturing regime change in models with unit roots.

Table 1 below presents various quantiles of the distribution introduced in Proposition 1 across alternative magnitudes of λ_0 the trimming parameter. The values have been obtained via standard simulations under a unit root DGP with $NID(0,1)$ errors and using $T = 2000$ across $N = 2000$ replications.

Table 1. Quantiles of the Limiting Distribution of $SupWaldA$

λ_0	0.50	0.90	0.95	0.975	0.99
0.05	13.74	20.77	23.34	25.41	28.77
0.10	13.14	20.20	22.77	24.78	27.89
0.15	12.61	19.61	21.87	24.18	26.45

3 Finite Sample Size and Local Power Considerations

Using $\Delta y_t = e_t$ as our DGP Table 2 below presents empirical size estimates using $\lambda_0 = 0.10$ and the above cutoffs.

Table 2. Empirical Size Estimates of $SupWaldA$

$Nominal$	2.5%	5.0%	10.0%
$T = 200$	3.10	6.00	11.60
$T = 400$	2.97	5.10	10.50
$T = 800$	2.40	4.25	10.25

The test displays a slight tendency to overreject under a 2.5% nominal size but is otherwise accurate across all scenarios.

Next, we are interested in scenarios whereby $\rho_1 = \rho_2 = c/T$ for $c < 0$ while the parameters associated with the deterministic components are kept time invariant. This scenario corresponds to a linear local to unit root model. Letting $DF_{\tau,\infty}(c)$ denote the limiting distribution of the t ratio for testing $\rho = 0$ in $\Delta y_t = \mu + \delta t + \rho y_{t-1} + e_t$ when $\Delta y_t = (c/T)y_{t-1} + e_t$ and whose expression is given under Theorem 3(d) in Phillips and Perron (1988, p. 342) we have the following result.

Proposition 2. *Under the same assumptions as in Caner and Hansen (2001),* $\theta_1 = \theta_2, \rho_1 = \rho_2 = c/T$ *and as* $T \to \infty$ *we have* $\sup_\lambda W_T^A(\lambda) \Rightarrow \sup_\lambda BB(\lambda)/\lambda(1-\lambda) + DF_{\tau,\infty}(c)^2$.

The above result illustrates the local power properties of our test statistic under linearity but with a local to unit root process. It is interesting to note that the first component of the limiting distribution remains unaffected by whether $\rho_1 = \rho_2 = 0$ or $\rho_1 = \rho_2 = c/T$. Our finite sample based power experiments are geared towards uncovering departures from the linear unit root $\rho_1 = \rho_2 = 0$ while maintaining $\theta_1 = \theta_2$. Our DGP is $\Delta y_t = \mu + (c/T)y_{t-1} + e_t$ and Table 3 below presents the estimated correct decision frequencies using our asymptotic quantiles.

Table 3. Power Properties of $SupWaldA$

c	-1	-10	-15	-20	-25	-30	-35	-40	-50
$T = 200$	3.30	10.60	19.10	32.70	52.10	69.50	84.60	93.50	99.50

Clearly power increases towards one as we move away from the unit root but is typically low for values of c up to around -30 which corresponds to an autoregressive parameter of 0.85. Beyond such magnitudes power quickly reaches 100%. This is very much in line with the the power properties of traditional unit root tests (see for instance Table 1 in Phillips and Perron (1988)).

It is also interesting to explore the behaviour of $SupWaldA$ when deviations occur in one direction from the null. We use Δy_{t-1} as our threshold variable and set $\gamma = 0$ as the corresponding true threshold parameter i.e. $\Delta y_t = (c/T)y_{t-1}I(\Delta y_{t-1} > 0) + e_t$ (case (i) say) while in the second scenario $\Delta y_t = (c/T)y_{t-1}I(\Delta y_{t-1} <= 0) + e_t$ (case (ii)). Empirical rejection frequencies are displayed in Table 4 below and suggest that the properties are unaffected by whether the exact unit root is present in the first regime or the second one.

Table 4. Further Power Properties of $SupWaldA$

c	-1	-10	-15	-20	-25	-30	-35	-40	-50
(i)	3.70	9.40	16.60	27.30	40.10	55.80	69.40	81.00	94.10
(ii)	3.90	10.60	18.90	29.80	41.70	58.40	73.10	82.90	96.00

4 Conclusions & Extensions

We have proposed a test of the joint null of linearity and a unit root within a TAR(1) model with deterministic components. One obvious limitation of our approach is our focus on a first order autoregression which ruled out the inclusion of lagged dependent variables. It is straightforward to establish that our results continue to hold if our model in (1) is augmented with lagged dependent regressors *provided that* their associated parameters are assumed to be time invariant as for instance in $\Delta y_t = (\theta_1' w_{t-1} + \rho_1 y_{t-1})I_{1t-1} + (\theta_2' w_{t-1} + \rho_2 y_{t-1})I_{2t-1} + \sum_{j=1}^{k} \psi_j \Delta y_{t-j} + \tilde{e}_t$ and are also excluded from our earlier restriction matrices (i.e. the parameters associated with the lagged dependent variables must assumed to be time invariant).

REFERENCES

Caner, M., Hansen, B. E., 2001. Threshold Autoregression with a Unit Root. Econometrica. 69, 1555-1596.

Hamilton, J. D., 1994. Time Series Analysis. Princeton University Press.

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

Phillips, P. C. B., Perron, P., 1988. Testing for a Unit Root in Time Series Regression. Biometrika, 75, 335-346.

Pitarakis, J., 2008. Threshold Autoregressions with a unit root: Comment. Econometrica, Vol. 76, pp. 1207-1217.

Pitarakis, J., 2011. Joint Detection of Structural Change and Nonstationarity in Autoregressions. Discussion Papers in Economics and Econometrics No. 1102, University of Southampton.

APPENDIX

PROOF OF PROPOSITION 1: With $D_T = diag(\sqrt{T}, T^{3/2}, T)$ and using Theorem 3 in Caner and Hansen (2001) and Lemma 3.1 in Phillips (1988) we have $D_T^{-1} X_1{}' X_1 D_T^{-1} \Rightarrow \lambda \int_0^1 \overline{B}(r)\overline{B}(r)'$ with $\overline{B}(r) = (1, r, B(r))$. Similarly, $D_T^{-1} X_2{}' X_2 D_T^{-1} \Rightarrow (1-\lambda) \int_0^1 \overline{B}(r)\overline{B}(r)'$ and $D_T^{-1} X' X D_T^{-1} \Rightarrow \int_0^1 \overline{B}(r)\overline{B}(r)'$. Next, setting $\sigma_e^2 = 1$ for simplicity and no loss of generality and using Theorem 2 in Caner and Hansen (2001), we have $D_T^{-1} X_1{}' e \Rightarrow \int \overline{B}(r) dB(r, \lambda)$ and $D_T^{-1} X' e \Rightarrow \int \overline{B}(r) dB(r)$. At this stage it is also convenient to note that we can write $D_T^{-1} X_1{}' e - \lambda D_T^{-1} X' e \Rightarrow \int_0^1 \overline{B}(r) dG(r, \lambda)$ where $G(r, \lambda) = B(r, \lambda) - \lambda B(r, 1)$ is known as a Kiefer process. We note that this latter random variable is mixed normal with variance $\lambda(1-\lambda)$ due to the independence of $G(r, \lambda)$ and $\overline{B}(r)$ since $E[G(r_1, \lambda_1) B(r_2, 1)] = 0$ and both processes are Gaussian. Finally, it is convenient to note the algebraic identity $W_T^A(\lambda) \equiv (u'V(V'V)^{-1} R_B' [R_B (V'V)^{-1} R_B']^{-1} R_B (V'V)^{-1} V'u + u'X(X'X)^{-1} R_L' [R_L (X'X)^{-1} R_L']^{-1} R_L (X'X)^{-1} X'u)/\hat{\sigma}^2$ where $R_L = (0\ 0\ 1)$ and $R_B = (I_3 \quad -I_3)$. The result now follows through the use of the CMT and the above intermediate limiting results.

PROOF OF PROPOSITION 2. The proof of Proposition 2 follows identical lines to our proof of Proposition 1 and is therefore omitted.